Lorraine T. Benuto
Brian D. Leany *Editors*

Guide to Psychological Assessment with African Americans

Springer

Guide to Psychological Assessment with African Americans

Lorraine T. Benuto • Brian D. Leany
Editors

Guide to Psychological Assessment with African Americans

Springer

Editors
Lorraine T. Benuto
Administrative Faculty, Department of Psychology
University of Nevada
Reno
Nevada
USA

Northcentral University
Core Faculty in the Graduate School
Prescott Valley
Nevada
USA

Brian D. Leany
Lake's Crossing Center
Maximum Security Facility for Forensic Mental Health Services
Sparks
Nevada
USA

Northcentral University
Core Faculty in the Graduate School
Prescott Valley
Nevada
USA

ISBN 978-1-4939-1003-8 ISBN 978-1-4939-1004-5 (eBook)
DOI 10.1007/978-1-4939-1004-5
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014955077

© Springer Science+Business Media New York 2015
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my Personal Penguins: Brian, Jorge, and Jack.

—L.T.B.

To my parents Jackie and Doug, who always described college and hard work as if they were as intrinsic and beneficial to life, as the rising and setting of the sun.

—B.D.L.

Also to Janice Stern, our editor, for your unwavering support and patience.

—L.T.B and B.D.L

Contents

Contributors

Tahirah Abdullah Department of Psychology, University of Massachusetts Boston, Boston, MA, USA

Daniel Allen Department of Psychology, University of Nevada, Las Vegas, Las Vegas, NV, USA

Leyla Bahojb Nouri Center for Mental Health Disparities, Psychological & Brain Sciences, University of Louisville, Louisville, KY, USA

Samantha M. Behbahani Psychology Department, Carlos Albizu University, Miami, FL, USA

Natalie Bennett Administrative Faculty, Department of Psychology MS0296, University of Nevada, Reno, NV, USA

Lorraine T. Benuto Administrative Faculty, Department of Psychology MS0296, University of Nevada, Reno, NV, USA

Northcentral University, Core Faculty in the Graduate School, Prescott Valley, AZ, USA

Alfiee M. Breland-Noble Department of Psychiatry, Research Division, Georgetown University Medical Center, Washington DC, USA

Tamara L. Brown College of Juvenile Justice and Psychology, Prairie View A & M University, Prairie View, TX, USA

F. Antoinette Burriss Department of Social Work, North Carolina State University, Raleigh, NC, USA

Xavier E. Cagigas Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA

L. Kevin Chapman Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY, USA

A. J. Culbreth Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA

Ryan C. T. DeLapp Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY, USA

Brad Donohue Psychology Department, University of Nevada, Las Vegas, NV, USA

Erin Dowdy Department of Counseling, Clinical, and School Psychology, University of California Santa Barbara, Santa Barbara, CA, USA

Jena Garric University of Nevada, Reno, NV, USA

Ashly L. Gaskin Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Robert M. Greenstone Psychology Department, Carlos Albizu University, Miami, FL, USA

Shanette M. Harris Department of Psychology, University of Rhode Island, Kingston, RI, USA

Shawn C.T. Jones Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Allison Kalpakci Department of Psychology, University of Houston, Houston, TX, USA

Brian D. Leany Lake's Crossing Center, Maximum Security Facility for Forensic Mental Health Services, Sparks, NV, USA

Northcentral University, Core Faculty in the Graduate School, Prescott Valley, AZ, USA

Daniel B. Lee Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Emily Malcoun Philadelphia, PA, USA

M. Marzell Family Research & Services, University of Nevada, Las Vegas, NV, USA

Lindsay F. Morra Department of Psychology, State University of New York, Binghamton, NY, USA

Enrique W. Neblett Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Marc A. Norman University of California, San Diego, CA, USA

Tara C. Raines Department of Educational Psychology and Higher Education, University of Nevada Las Vegas, Las Vegas, NV, USA

Allyn Richards Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY, USA

Broderick Sawyer Center for Mental Health Disparities, University of Louisville, Psychological & Brain Sciences, Louisville, KY, USA

Carla Sharp Department of Psychology, University of Houston, Houston, TX, USA

Cyndy G. Soto Administrative Faculty, Department of Psychology MS0296, University of Nevada, Reno, NV, USA

Jason Sotomayor Department of Psychology, Georgetown University, Washington DC, USA

G. P. Strauss Psychology Department, State University of New York at Binghamton, Binghamton, NY, USA

Nicholas S. Thaler Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA

April D. Thames Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, USA

Jennifer Twyford raduate School of Psychology, California Lutheran University, Oxnard, CA, USA

Jessica Urgelles Department of Psychology, University of Nevada, Las Vegas, NV, USA

Ebony S. Vinson Department of Psychology, University of Kentucky, Lexington, KY, USA

Chad T. Wetterneck Rogers Memorial Hospital, Oconomowoc, WI, USA

Monnica T. Williams Center for Mental Health Disparities, University of Louisville, Psychological & Brain Sciences, Louisville, KY, USA

Monnica T. Williams Center for Mental Health Disparities, Psychological & Brain Sciences, University of Louisville, Louisville, KY, USA

Davor Zink Department of Psychology, University of Nevada, Las Vegas, Las Vegas, NV, USA

1

Conducting Assessments and Psychological Evaluations with African American Clients

Lorraine T. Benuto and Brian D. Leany

Introduction to Assessment with African American Clients

The Cultural Sensitive Movement (CSM) emerged nearly 40 years ago (Sue 1977a, b; Constantine and Sue 2005) and ample scholarly activity on the topic of cultural sensitivity has materialized since. In fact, in 1990 the American Psychological Association (APA) developed cultural competency guidelines and then revised them in 2002. While all six guidelines are relevant to this book, Guidelines 2 and 5 are particularly relevant here. Guideline 2 is as follows: *Psychologists are encouraged to recognize the importance of multicultural sensitivity/responsiveness, knowledge, and understanding about ethnically and racially different individuals*. In this book we provide an overview of African American cultural characteristics that are relevant to psychological assessment and evaluations. We also provide specific guidelines for clinicians and researchers who work with African American clients and this aligns with Guideline 5. Guideline 5 specifies that *Psychologists strive to apply culturally–appropriate skills in clinical and other applied psychological practices*. This guideline combined with the mental health field's shift towards evidence-based practice and stance that expert opinions should be grounded in science and based upon methodologically sound data (Lilienfeld and Landfield 2008) make it clear that specific guidelines that are rooted in science and theory for identified cultural groups are highly necessary.

As indicated above, the mental health field is increasingly oriented toward evidence-based practice (Lilienfeld and O'Donohue 2007; Ollendick 2012) not only with regard to treatment of psychological disorders but also with regard to assessment. In fact Hunsley and Mash (2007) specified that evidence-based assessment (EBA) emphasizes the use of research and theory to guide the selection of assessment domains, the manner in which the assessment is administered, the measures used, and the assessment process. These authors also called for the development of EBA guidelines, emphasized the importance of diversity, and called for dissemination of accurate and current information on EBAs. The crux of this book is the timely development of empirically based (and where empirical evidence is absent theoretically derived) recommendations for assessments with African Americans.

L. T. Benuto (✉)
Administrative Faculty, Department of Psychology
MS0296 University of Nevada, Reno, NV 89557, USA
e-mail: dr.benuto@gmail.com

B. D. Leany · L. T. Benuto
Northcentral University, Core Faculty in the Graduate School, 10000 E University Drive, Prescott Valley, AZ 86314, USA
e-mail: leanyb@gmail.com

B. D. Leany
Lake's Crossing Center, Maximum Security Facility for Forensic Mental Health Services, 500 Galletti Way, Sparks, NV 89504, USA

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_1,

Organization of the Book

In this book we called upon experts to identify the range of psychological assessments administered by clinicians and to discuss these measures as they pertain to the African American client. This book contains chapters that emphasize specific assessment domains (e.g., personality, self-report, IQ, etc.); where appropriate chapters relevant to a certain types of psychological assessment (e.g., forensic, neuropsychological, school-based, etc.); and chapters specific to many of the major diagnostic categories (e.g., eating disorders, mood disorders, etc.) in the Diagnostic and Statistical Manual of Mental Disorders IV-Text Revision (DSM-IV-TR: APA 2000). As we were preparing this book the DSM-5 came to press. However, we opted to focus on the DSM-IV-TR as only now that the DSM-5 has been released, will research on assessment measures begin to be conducted on measures that align with diagnostic categories housed within the DSM-5.

Each chapter contains a general overview of cultural considerations that should be made with African American clients as they relate to the diagnostic category in question. An overview of the most commonly used and/or frequently researched assessment measures and the research that has been conducted on these measures with African American participants is also included in each chapter. It is important to note that the African American population merits a number of specific considerations. These are detailed within the remained of this chapter and in Chap. 2. In deed the ability to identify experts who felt they had sufficient empirical evidence to present recommended considerations for African Americans was not always possible. As an example of this difficulty, the original proposal for this book contained chapters related to the assessment of sexual dysfunction and the use of projective assessments with African Americans, but authors solicited for these chapters cited a lack of sufficient empirical evidence for these domains and thus these chapters were ultimately omitted from this book.

Cultural Considerations When Working with African Americans

In addition to the assessment-specific chapters contained in this book, Chap. 2 (Brown et al. 2014) offers an overview of the cross-cultural considerations that should be made when conducting psychological assessments or evaluations with African American clients. While this section of this introductory chapter is not intended to provide a thorough overview of these considerations, we felt it necessary here at the outset to highlight the unique nature of this group and its relevance to psychological assessment. Before we proceed, it is important to discuss the use of terminology frequently used in this book. While the terms "Black " and "African American" often appears to be used interchangeably in this text, where possible we used terminology consistent with the original scholarly source from which the content we are discussing was pulled. Typically, the terms "Black" and "African American are used to distinguish persons of African descent who have immigrated to the USA (i.e., they were not born in the USA) from those who were born in the USA (i.e., African Americans: Brown et al. 2014).

According to census.gov (2013) African Americans constituted 13.1 % of the population in 2012. Despite that this group represents a substantial part of the US cultural panorama, it is clear that this group has been largely underrepresented in the literature (Coley and Barton 2004). This may be due to distrust of the medical/scientific community, poor access to care, poor recruiting strategies, etc. (Shavers-Hornaday et al. 1997). Indeed in providing culturally sensitive services to African American, it is important to possess background knowledge of the historical events that have shaped the attitudes toward health and healthcare providers that predominate among this group (for a detailed account of such historical events, see Hargrave 2010). Specifically African Americans may be suspicious of and reluctant to trust healthcare providers because of experiences with prejudice and discrimination by healthcare providers. This mistrust may due to misdiagnosis and clinical bias both of which occur frequently in the psychiatric assessment and development of treatment plans for African American (Hargrave 2010). Thus the assessing clinician may wish to ensure that rapport has been established and keep in mind that test-taking attitudes can impact performance and response style.

Using Theory and Research to Guide Assessment

As indicated at the outset of this chapter Hunsley and Mash (2007) specified that EBA emphasizes the use of research and theory to guide the assessment process. This aligns well with overarching goal of this book, which was to provide EBA recommendations for use with African American clients. It is important to note however that in many cases (depending on the assessment domain or disorder of interest) research on this population was sparse. For example, a review of 60 years of personality assessment research demonstrated that only 6% of participants were identified as ethnic or racial minorities (Holaday and Boucher 1999). Despite that this review is over 10 years old, given the population of African Americans in the USA and that the cultural sensitivity movement emerged nearly 40 years ago, it is discerning that ethnic minorities have not been included in the research at a higher rate. Thus, where data is absent we turned to theory to guide the development of appropriate guidelines.

In sum, while not ideal, at times clinicians and researchers may need to use their knowledge of African American cultural characteristics and hypothesize how these cultural characteristics may influence symptom presentation and/or performance on specific measure. These hypotheses can then be used to guide the interpretation of test results and/or recommendations based on an assessment. Within this text where direct recommendations based on empirical evaluations of specific measures cannot be made (due to an absence of empirical literature), a theoretical position may be taken and African American cultural values may be considered in light of the construct being assessed. While in general clinicians and researchers should avoid making recommendations based on anthropological assertions that have not been empirically investigated, for measures that have high face validity this may be a reasonable approach to take when empirical data is absent (Leany and Benuto In Press).

Racial and Ethnic Concordance

Racial and ethnic concordance refers to concordance between a client/patient and the healthcare provider. A primary focus of the research on concordance is related to perceived quality of care and its relationship to utilization (an exemplar of this research can be seen in Saha et al. 1999). Saha, et al. reported better outcomes when the doctor–patient race/ethnicity was concordant, and also reported a better utilization of health services (e.g., preventative care). As psychologists (even those not of a Gestalt or CBT orientation), we can hopefully appreciate the impact of this perception on behavior (as well as a measure of cognition). In addition to the emphasis on patient perceptions, when searching outcome data as it relates to the concordance between patient and clinician ethnicity and race (even when utilizing a key term of psychologist rather than physician) results were heavily weighted (almost exclusively) towards medical domains. Because of these perceptions and corresponding improvements, a primary concern heralded in the literature is the shortage of physicians who are of the same race or ethnicity as the clients they serve (Komaromy et al. 1996), and a resulting overrepresentation of same race/ethnicity patients on those physicians caseload as well as the number of those patients who utilize government subsidized health insurance (e.g., Medicaid and Medicare). This presents a challenge not only for the medical field, but if one assumes this trend holds true for mental health clinicians, there is also a disparity between patients and available clinicians that identify as African American (Komaromy et al. 1996). Indeed a number of studies identified subtle biases that may negatively impact rapport (Dovidio et al. 2008). Ultimately Dovidio et al.'s evaluation of the empirical literature revealed that there was a greater occurrence of longer visits for race concordant contacts, but disproportionately less conversation for the clinician as well as less encouragement for patient involvement in treatment decision making (specifically when the clinician was White and the patient Black) when racial concordance was absent, with the latter condition yielding less information provided to the clinician, and poorer outcomes (e.g., less utilization of preventative services and poor follow-up care).

While the clinician cannot change his or her own race or ethnicity to meet client needs, clinicians can educate themselves by increasing their expertise regarding the cultural characteristics of the racial and ethnic groups with whom they interact (the fact that you have picked up this book suggests a willingness to take that step). This is in alignment with Guideline 2 of the multicultural guidelines put forth by the APA, which specifies that psychologists should recognize the importance of being culturally sensitive and understanding ethnically and racially difference individuals.

Further, there are common factors that relevant to any assessment regardless of the ethnicity or race of the client. For example, Mulvaney-Day et al. (2011) provided examples and context for some African American patients' expectations of a mental health clinician. Their research identified patient themes that emphasized listening, understanding and a desire for the clinician to find a way to "bridge the gap" for obvious differences (described as not only racial differences, but those of socioeconomic status, e.g., education). The qualitative reports from the Mulvaney-Day et al. (2011) study suggested a less directive approach was desired and that, while those patients in the study reported a respect for the professional's opinion. In particular, when the clinician was not of the same race as the patient, the patient had greater expectations that they be appropriately seen as an individual and be allowed to describe their unique circumstances or position. Thus, the aforementioned research suggests that, in particular, when there is a racial discrepancy between the patient and clinician, a clinician must be willing to take the additional time to not just collect information for objective assessment, but also listen to what may seem like ancillary information in order to "bridge" the gap between patient and clinician. The clinician must also recognize that there is evidence of implicit behaviors that can result in suboptimal doctor–patient relationships, and thus be more cognizant of this potential in order to be purposeful in efforts to reduce this potential (e.g., checking for understanding of recommendations, encouraging of patient decision-making, avoiding cutting short the evaluation, etc.).

Education and Economic Status Count: Using Adjusted Norms

As illustrated throughout this text, disparity in performance on psychological assessments has been demonstrated across multiple domains. While, the field of psychology has a notorious history of biased assessment for intellectual assessment of African Americans (Anglin and Kwate 2009) a number of current iterations of psychological assessment instruments have made purposeful efforts to include samples that are more representative of the US population and that incorporate a broader range of educational and economically diverse groups across race and ethnicity.

The chapters on IQ and Neuropsychological assessment represent the bulk of the discussion of differential performance for African Americans, in particular as compared to their Caucasian counterparts. Many of the works identified in those chapters describe a reduction of bias that occurs when one uses education and economic status based norms (Heaton et al. 2003Taylor and Heaton 2001). These norms do much to improve psychometric properties for measures of intellect and neurocognitive abilities, but there is still emerging data that suggests we could do more to account for additional, more subtle variables such as the quality of education (Manly et al.

2002; Manly et al. 2004; O'Neill 1990). It should be noted, that specific to neuropsychological assessment, the method of assessment for premorbid intellectual functioning can be problematic. Research for this domain has shown problems of disparity between reported educational attainment and comparable performance on achievement-based measures of premorbid functioning (Baker et al. 1996; Manly et al. 2002). While many of the personality based measures and forensic tools describe comparable psychometric data, when using age and education as well as race-based norms, there are the aforementioned limitations for intellectual and neuropsychological assessment. Thus, it is important for the clinician to be aware of the availability and limitations of normative comparison groups, in particular for these two domains, as the consequences are much higher than they would be for a diagnosis of depression or an anxiety based disorder. Consequently, relevant chapters discuss the limitations of various measures and their available norms as well as provide recommendations for the clinician.

Clinician/Rater Bias

In addition to problems related to normative data, there is also literature that documenting clinician and rater bias, unrelated to choice of norms. This data suggests that there is a tendency to overdiagnose African Americans with psychotic disorders, including schizophrenia, usually attributed to a misattribution of behaviors (e.g., spirituality and religiosity) as symptoms of psychosis (Neighbors et al. 2003), and that this is most likely to occur in hospital settings (Neighbors et al. 1999; Neighbors 2003) as well as when a clinician does not rely on a standardized measure of assessment (e.g., a structured or semistructured interview; First et al. 2002). In addition to a misattribution of symptoms by the clinician, the culture may communicate information that is not specifically clinician or rater bias, but a phenomenon that may create a response bias by the examinee.

Researchers have reported a "stereotype" effect in which the participants performance is unduly influenced by societal influences of both overt and covert racism, that result in the examinee responding in a manner that supports these stereotypical values (e.g., decreased performance on achievement and academic based assessments that rely heavily on performance for tasks that are most salient for those of the dominant culture; Steele and Aronson 1995; Thames et al. 2013). It should be noted that these effects are not race specific, in that Steele (1997) was able to demonstrate a similar effect for White males. Further, as forensic based evaluations (e.g., competency and capacity) rely on mental health diagnoses and intellectual functioning, it too is impacted by these biases, with even higher stakes for misattribution of symptoms or diagnostics error.

In these instances, the recommendation is less straight-forward than a simple suggestion to apply race and education based norms, but, consistent with the aforementioned APA guidelines 2 and 5 discussed at the outset of this chapter, is rather a reminder to the clinician to be vigilant in their awareness of relevant literature. Further, the recommendations strongly encourage a direct inquiry about perceptions of racial equality and discrimination. Specific to the differentiating between potential symptoms of psychosis and faith or spirituality, in addition to the use of structured or semistructured interviews (like the SCID), the clinician should make direct inquiries about an individual's spirituality and religion, and greater consideration to these factors should be made, when stakes are higher, such as for forensic evaluations.

Summary and Conclusions: Trends and Themes in EBA with African Americans

Through the editing (and authoring) process a number of trends and themes were observed. The first is that individuals who work with African Americans should at the very least have background knowledge of the historical events that have shaped the culturally derived attitudes toward health and healthcare providers. Indeed, African Americans may present with distrust of healthcare providers and exhibit a preference for racial concordance. This leads us to our second point, which is that there is ample evidence that clinicians can act in a biased manner when administering and interpreting test results. Thus practitioners should be aware of their own biases and find appropriate means to resolve them. The third point we make herein is that clinicians and researchers alike should strive to use empirical evidence where possible to guide the choices they make in terms of assessment choices, the use of norms (rather age-adjusted, education adjusted, or norms that are specific to African Americans), and where not possible clinicians may choose to use theoretical information to aid in their decisions. Lastly, as with any assessment no single test should be used to make any decision. In fact, multiple sources of information should be gathered. Assessment should occur in a hypotheses-testing context whereby the clinician administers test and collects data (e.g., through record review, collateral contacts, etc.) to support or refute test findings. Along these lines, clinicians should ensure that assessment practices examine the full-range of diagnostic criteria, with particular care to assess for symptoms that may be traditionally underreported in African American samples (e.g., cognitive symptoms of anxiety

As evidenced above and throughout this book, it is clear that clinicians and researchers must make cultural considerations when conducting psychological assessments or psychological evaluations with African American clients. However, there remain advancements to be made in terms of EBAs for African Americans. Thus, in alignment with Guideline 4 of the APA's cultural competency guidelines we recommend that: *Culturally sensitive psychological researchers are encouraged to recognize the importance of conducting culture–centered and ethical psychological research among persons from ethnic, linguistic, and racial minority backgrounds*.

References

American Psychological Association. (1990). *Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

American Psychological Association. (2002). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. http://www.apa.org/pi/multiculturalguidelines/homepage.html. Accessed 22 Dec 2007.

Anglin, D. M., & Kwate, N. A. (2009). Social and cultural factors in the cognitive and clinical assessment of African American adults. In H. A. Neville, B. M. Tynes, & S. O. Utsey (Eds.), *Handbook of African American psychology* (pp. 459–468). Thousand Oaks: Sage Publications. Inc

Baker, F. M., Johnson, J. T., Velli, S. A., & Wiley, C. (1996). Congruence between education and reading levels of older persons. *Psychiatric Services, 47*(2), 194–196.

Brown, T. L., Vinson, E. S., & Abdullah, T. (2014). Cross-cultural considerations with African American clients: A perspective on psychological assessment. In L.T. Benuto & B. D. Leany (Eds.), *Guide to psychological assessment with African Americans*. New York: Springer.

Coley, R. J., & Barton, P. E. (2004). The chronic underrepresentation of African Americans in medicine. *ETS Policy Notes, 12,* 1–8.

Constantine, M. G., & Sue, D. W. (Eds.). (2005). *Strategies for building multicultural competence in mental health and educational settings*. Hoboken: Wiley.

Dovidio, J. F., Penner, L. A., Albrecht, T. L., Norton, W. E., Gaertner, S. L., & Shelton, J. (2008). Disparities, and distrust: The implications of psychological processes for understanding racial disparities in health and health care. Social Science & Medicine, 67(3), 478–486. doi:10.1016/j.socscimed.2008.03.019.

First, M. B., Bell, C. C., Cuthbert, B., Krystal, J. H., Malison, R., Offord, D. R., et al. (2002). Personality disorders and relational disorders: A research agenda for addressing crucial gaps in DSM. In D. J. Kupfer, M. B. First, & D. A. Regier (Eds.), *A research agenda for DSM-V* (pp. 123–199). Arlington: American Psychiatric Association.

Hargrave, R. (2010). Health and health care of African American older adults. http://geriatrics.stanford.edu/ethnomed/african_american/downloads/african_american.pdf. Accessed Mar 2014.

Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Demographic effects and use of demographically corrected norms with the WAIS-III and WMS-III. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, & R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 181–210). San Diego: Academic. doi:10.1016/B978-012703570-3/50010-9.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994a). An exchange: Part I: Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher,* 23(3), 5–14.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994b). Does money matter?: A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher,* 23(3), 5–14. (Reprinted as pp 63–84 in J. Grao and A. Ipia (Eds.), Economa de la Education. Vitorio-Gasteiz: Servicio Central de Publicationes del Gobierno Vasco, 1966).

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994c). Money does matter somewhere: A reply to Hanushek. *Educational Researcher, 23*(4), 9–10.

Holaday, M., & Boucher, M. (1999). Journal of personality assessment: 60 years. *Journal of Personality Assessment*, 72(1), 111–124. doi:10.1207/s15327752jpa7201_7.

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology, 3,* 29–51.

Leany, B., & Benuto, L. T. (In Press). Forensic assessment with Asian clients. In L. T. Benuto & N. Thaler (Eds.) *Guide to psychological assessment with Asian clients*. New York: Springer.

Lilienfeld, S. O., & Landfield, K. (2008). Issues in diagnosis: Categorical vs. dimensional. In W. Craighead, D. J. Miklowitz, & L. W. Craighead (Eds.), *Psychopathology: History, diagnosis, and empirical foundations* (pp. 1–33). Hoboken: Wiley.

Lilienfeld, S. O., & O'Donohue, W. T. (2007). *The great ideas of clinical science: 17 principles that every mental health professional should understand*. New York: Routledge.

Komaromy, M., Rumbach, K. G., Drake, M., Vranizan, K., Lurie, M., et al. (1996). The role of black and Hispanic physicians in providing health care for underdeserved population. *New England Journal of Medicine, 334,* 1305–1310.

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society, 8*(3), 341–348. doi:10.1017/S1355617702813157.

Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology, 11*(1), 37–46.

Mulvaney-Day, N., Earl, T., Diaz-Linhart, Y., & Alegria, M. (2011). Preferences for relational style with mental health clinicians: A qualitative comparison of African American, Latino and non-Latino White patients. *Journal of Clinical Psychology, 67*(1), 31–44.

Neighbors, H. W., Trierweiler, S. J., Munday, C., Thompson, E. E., Jackson, J. S., Binion, V. J., & Gomez, J. (1999). Psychiatric diagnosis of African Americans: Diagnostic divergence in clinician structured and semistructured interviewing conditions. *Journal of the National Medical Association, 91,* 601–612.

Neighbors, H. W., Trierweiler, S. J., Ford, B., & Muroff, J. R. (2003). Racial differences in diagnosis using a semi-structured instrument: The importance of clinical judgement. *Journal of Health and Social Behavior, 44,* 237–256.

Ollendick, T. H. (2012). The role of assessment and case conceptualization in evidence-based practice. Psyccritiques, 57(9). doi:10.1037/a0027211.

O'Neill, J. (1990). The role of human capital in earnings differences between black and white men. *The Journal of Economic Perspectives, 4*(4), 25–45.

Saha, S., Komaromy, M., Koepsell, T. D., & Bindman, A. (1999). Patient-physician racial concordance and the perceived quality and use of health care. *Archives of Internal Medicine, 159,* 997–1004.

Shavers-Hornaday, V. L., Lynch, C. F., Burmeister, L. F., & Torner, J. C. (1997). Why are African Americans under-represented in medical research studies? Impediments to participation. *Ethnicity & Health, 2,* 31–45.

Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist,* 52(6), 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811. doi:10.1037/0022-3514.69.5.797.

Sue, S. (1977a). Community mental health services to minority groups: Some optimism, some pessimism. *American Psychologist, 32*(8), 616–624. doi:10.1037/0003-066X.32.8.616.

Sue, D. W. (1977b). Counseling the culturally different: A conceptual analysis. *Personnel & Guidance Journal, 55*(7), 422–425. doi:10.1002/j.2164-4918.1977.tb05281.x.

Taylor, M. J., & Heaton, R. K. (2001). Sensitivity and specificity of WAIS-III/WMS-III demographically corrected factor scores in neuropsychological assessment. *Journal of the International Neuropsychological Society, 7*(7), 867–874.

Thames, A. D., Hinkin, C. H., Byrd, D. A., Bilder, R. M., Duff, K. J., Mindt, M., et al. (2013). Effects of stereotype threat, perceived discrimination, and examiner race on neuropsychological performance: Simple as Black and White? *Journal of the International Neuropsychological Society, 19*(5), 583–593. doi:10.1017/S1355617713000076.

2

Cross-Cultural Considerations with African American Clients: A Perspective on Psychological Assessment

Tamara L. Brown, Ebony S. Vinson and Tahirah Abdullah

Cross-cultural assessment of psychological phenomena is not a new endeavor. It has its roots in early anthropological studies that attempted to understand the nature of people through the study of different cultures (Butcher and Garcia 1978). For anthropologists, a critical issue is figuring out how to ensure that, in the process of their investigations, they do not wittingly or unwittingly distort their field data to conform to preconceived expectations that are based on their own ethnocentric biases. This same issue confronts mental health professionals who seek to assess clients who are culturally different from them. Stated differently, if the goal of assessment is to learn something about another person, how do we ensure that the approaches and instruments we use allow us to clearly and accurately understand the clients worldview and level of functioning rather than merely provide a reflection of our own worldview, biases, and expectations? To the extent that test takers select instruments that are valid in their own cultures but have not been shown to be valid in the client's culture, the risk of clinical errors such as interpreting difference as deficit, over-pathologizing normality, overlooking symptoms, and misdiagnosis is magnified. These risks of clinical errors apply to the assessment of African Americans because, despite researchers' and psychologists' assumptions of homogeneity, the label "African American" includes people from many different national, linguistic, ethnic, racial, cultural, and social backgrounds.

Due to the African slave trade, African Americans have been a major part of the US population since the country's founding, accounting for nearly 20 % of the US population counted in the 1790 census (Kent 2007). Though the ending of the slave trade in 1808 considerably reduced the flow of Africans to the USA, changes in US immigration laws and technology, and societal unrest in African countries led to a seven-fold increase in the number of foreign-born Blacks between 1960 and 1980, and between 1985 and 2005 their numbers tripled. According to research compiled by Helina Faris of the Center for American Progress, an independent, nonpartisan educational institute, Black immigrants comprise 8 % of the US foreign-born population; more than half come from the Caribbean, the bulk of the rest come from northern and sub-Saharan Africa, and a small number come from Europe and Canada. Indeed Black immigrants account for more than one quarter of the Black population in New York, Boston, and Miami. Black immigrants enter the USA through multiple pathways. Most

---

T. L. Brown (✉)
College of Juvenile Justice and Psychology, Prairie View A & M University, P. O. Box 519, Prairie View, TX 77446, USA
e-mail: tlbrown@pvamu.edu

E. S. Vinson
Department of Psychology, University of Kentucky, 207a Kastle Hall, Lexington, KY 40506, USA
e-mail: ebony.vinson@uky.edu

T. Abdullah
Department of Psychology, University of Massachusetts Boston, McCormack Hall 04, Boston, MA 02125, USA
e-mail: Tahirah.Abdullah@umb.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_2, © Springer Science+Business Media New York 2015

(especially those from the Caribbean) arrive as legal permanent residents based on family ties; however, in 2009, 30 % of all Black African immigrants were refugees from Ethiopia, Somalia, Liberia, Sudan, and Eritrea, about 20 % entered through the diversity visa lottery program, and 400,000 were in the USA undocumented (Faris 2013).[1]

As with other immigrant groups, English is often not the first or primary language of Black immigrants. Although Caribbean-born Blacks are more likely to speak English at home than other immigrant groups, a fact that is not surprising when one considers that most are from former British territories and colonies, some Caribbean immigrants report speaking patois (an English-based dialect that combines English with West African languages), French/French Creole, or Spanish at home. African immigrants are likely to speak an African language at home (e.g., Amharic, Bantu, Kru, Swahili), though two thirds are also proficient in English. However, African immigrants who are fluent in English report experiencing problems being understood because their accents are unfamiliar, a problem that interferes with their school performance and their ability to find employment and obtain promotions (Kent 2007).

In terms of education, Black immigrants have more college education and higher rates of degree attainment than any other immigrant group in the USA, but because of discrimination they tend to earn low wages compared to similar educated workers, and, in 2011, had the highest unemployment rate of any foreign-born group in the USA (Faris 2013; Kent 2007). Black immigrants also tend to have traditional family arrangements. Census data from 2000 show that 76 % of African immigrant children and 65 % of Caribbean immigrant children lived in two-parent households, compared to 44 % of African American children. Although they endeavor to adapt to the US culture, many African and Caribbean immigrants also seek to hold onto their immigrant identity, in part to distinguish themselves from US-born Blacks. Many speak their languages of origin at home, live in ethnic enclaves, and maintain familial, social, and political ties to their countries of origin, and affiliate with US groups that provide a means of benefiting their particular ethnic communities. Most immigrant parents want their children to maintain the cultural values and heritage of their home countries; consequently, they teach their children their native language and culture, maintain their heritage, and send their children to spend time with relatives in their home country.

Like most immigrant groups, foreign-born Black immigrants gravitate toward metropolitan areas where the job opportunities are greater and where they find existing communities of US-born Blacks (Kent 2007). Indeed, Black immigrants are dispersed throughout the USA, especially in New York, California, Florida, New Jersey, and Texas. Kent (2007) reports that the top ten states where African and Caribbean immigrants live are similar except that the African list includes Virginia and Minnesota and the Caribbean list includes Florida and Connecticut.

From 2000 to 2010, the Black population in the USA grew by 15.4 % and this growth is expected to continue into the next decade (U.S. Census Bureau 2012). According to the 2010 US Census, self-identified Black people make up approximately 13.6 % of the US population. With the continued growth, it is imperative that psychologists acknowledge the potential role of ethnic/racial variation in their services and assessments. This chapter proposes that many commonly used psychological assessments may only be valid for the ethnic or cultural groups on whom they were developed. Considering the continued use of instrumentation developed by and standardized with European Americans, there are some important considerations and cultural issues that should be considered when assessing African Americans.

---

[1] Whenever we use the term "Black African immigrant" or similar terms, it is because this is the term used by the author of the published work we are citing. Generally, when authors use these terms it is to distinguish persons of African descent who have immigrated to the USA (i.e., they were not born in the USA) from those who were born in the USA (i.e., African Americans).

General Assessment Considerations

When conducting assessments with African Americans, the main issue that should be considered is validity. According to *Standards for Educational and Psychological Testing* (American Educational Research Association (AERA) 1999), validity is "the degree to which all the accumulated evidence supports the intended interpretation of test scores for the proposed purpose" (p. 11). Bias, which occurs when the validity of a test score differs for a group of individuals (Guerra and Jagers 1998), and fairness, the standard that individuals should be assessed equitably (AERA 1999), are often discussed in the context of testing with African Americans. Bias and fairness are issues of validity; that is, the presence of bias or absence of fairness would detrimentally impact the validity of an assessment. Validity and the many other issues subsumed within validity should be considered at each part of the assessment process: assessment instrument selection, administration, and interpretation of assessment results.

Assessment Selection Depending on the purpose of the assessment, inaccurate results can have significant implications for educational placement, diagnosis, or treatment interventions that could have a lasting impact on the client's life (Padilla 2001). Evidence that the assessment measure that is being used accurately characterizes the construct of interest (i.e., content validity evidence) is of significant relevance when selecting assessment instruments (Kaplan and Saccuzzo 2001). Many researchers agree that assessment instruments that have been "normed on majority group populations or developed using Eurocentric approaches cannot be indiscriminately used with individuals who differ from the normative population" (Padilla 2001, p. 5). Therefore, when selecting an assessment instrument to use, evidence related to the instrument's generalizability to African Americans should be considered. Awareness of whether or not norms for African Americans have been established for the instrument is also important.

Another important consideration is the language of the test. As Helms (1992) points out, White American English is considered to be the one and only standard that all Americans assessed for intelligence in particular must meet. For African Americans who speak "nonstandard" dialects of English, it has been perhaps incorrectly assumed that verbally-oriented intelligence scales are valid in assessing cognitive ability (Bracken and McCallum 2001). Some psychologists further claim that "traditional assessments of cognitive ability and written, oral language, and reading skills do not yield accurate results" (Gopaul-McNicol et al. 1998, p. 16). Additionally, with the growing population of Black immigrants from African and Caribbean countries, clinicians may encounter clients whose first language is not English.

Assessment Administration Issues of bias and fairness are essential to consider in assessment administration. There is an inherent power differential between clients and clinicians, with clinicians in the position of power (Ponterotto et al. 2001). This power differential may be intensified in a cross-cultural situation with an African American client and could introduce effects unrelated to the construct of interest that impact the assessment results.

Various examiner effects may impact clients, and could be sources of bias in assessments. The client's level of familiarity with the clinician may influence performance on assessments, particularly with clients of lower socioeconomic status (Skiba et al. 2002). Building and maintaining rapport is particularly important in cross-cultural assessments with African Americans. During assessments, clinicians' practice of maintaining a business-oriented, impersonal interaction style with African American clients may make clients feel uncomfortable or alienated (Dana 1996). Clients often do their own appraisal of clinicians to determine the clinician's level of caring, authenticity, and cultural understanding before becoming fully engaged in the assessment process. According to Dana (1996), if the client is content and comfortable following appraisal of the clinician, "there will be attempts to personalize the relationship. The outcome of these overtures can lead to a task orientation if sufficient personal regard has been established between client and assessor" (p. 476).

While examiner effects are important, there are also client effects that are imperative to be aware of and assess for when working with African Americans, specifically cultural mistrust and stereotype threat. Cultural mistrust has been defined as the "belief acquired by African Americans, due to past and ongoing mistreatment…, that Whites cannot be trusted" p. 299 (Terrell et al. 2009, p. 299). Although this mistrust may focus on particular people or groups of people, it could also generalize to institutions like the health care system. The extent to which the individual being tested trusts the clinician influences the client's engagement, cooperativeness, and effort in the assessment, which are factors that may impact accuracy of answers and performance on the assessment. Terrell et al. (1996) found that Black students with high levels of cultural mistrust scored lower on standardized intelligence tests. Similarly, there is some evidence that, especially for African American male college students, there is an interaction between examiner race and cultural mistrust in predicting scores on the Wechsler Adult Intelligence Scale—Revised (WAIS-R). Terrell et al. (1981) found that those who were high on cultural mistrust and were tested by a European American examiner had significantly lower WAIS-R scores than those who were high on cultural mistrust but were tested by an African American examiner. These findings may be related to minimal effort given by the test takers during evaluations because of the belief that the assessors held unfair views of them that would not change regardless of their effort or intelligence or a general distrust of the testing process. African Americans make up only a small percentage of mental health providers and assessors so cultural mistrust may be very important in understanding testing discomfort, specifically with testing that is not self-administered.

Stereotype threat, the risk of confirming negative stereotypes in situations where stereotypes related to a group one belongs to are made salient, can also influence test performance for African Americans (Steele and Aronson 1995). Fear of performing poorly that arises when stereotypes are elicited can create self-doubt and anxiety that have a detrimental impact on test performance. As a factor outside of the construct of interest that may influence the results of the assessment, stereotype threat presents the potential for bias.

Language can also be of concern in the administration of assessments. Mutual understanding between the provider and the client is necessary for accurate, effective assessment. It is especially important for providers to ensure that clients understand informed consent forms and all test instructions. It is also important for providers to understand the meaning of responses clients give during clinical interviews and other verbal forms of assessment. Paniagua (2005) provided an example of sentences consistent with "nonstandard" dialects spoken by some African Americans that have two different meanings, but may be easily misunderstood: "My child sick" and "My child be sick." While the first sentence indicates that the child is currently sick, the second sentence indicates that the child has an ongoing sickness. These responses would likely have different implications for the mother's stress level, and may result in different interpretations.

Assessment Interpretation During the assessment interpretation stage, the related issues of predictive validity and selection bias should be considered for African American clients. Often, assessments are used as evidence to predict some future outcome; however, clinicians should use particular caution in interpreting results in this way for African Americans. Selection bias occurs when test scores predict outcome criteria differently for two groups. For example, it has been suggested that low scores on the Scholastic Aptitude Test (SAT), a college entrance exam, do not predict college performance for African Americans (Franklin 2007). Also, a study by Whatley et al. (2003) examining racial identity and Minnesota Multiphasic Personality Inventory (MMPI) scores among Blacks has indicated that scores on the Immersion-Emersion subscale of the Racial Identity Attitude Scale-Black predict scores on MMPI scales four and nine. A 4–9 MMPI code type is generally associated with antisocial personality disorder features, which brings into question whether the MMPI may pathologize normal racial identity processes for African Americans (Whatley et al. 2003).

Cultural Issues

Culture is a broad term with multiple definitions but generally culture is defined as a social context where people share social norms, beliefs, values, language, and institutions (Guerra and Jagers 1998). Culture has been considered an unimportant, vague, and unsound construct that is often controlled for instead of studied (Jones and Rhee 2004) but core cultural values are represented in affect expression, behaviors, and cognitive styles, which may then influence the results of psychological assessment. For example, African Americans have been found to have less linear, more circular concepts of time, which may affect how they respond to timed tasks (Helms 1992). Social expectations and the environment may also influence symptom presentation, meaning attributions, and treatment seeking. "My nerves are bad!" and "My heart is heavy" are some examples of culturally influenced symptom presentations that may be observed among African Americans. Neighbors et al. (2003) suggested that racial differences in symptomology and emotional expression are uniquely challenging when implementing a categorical diagnostic system. Additionally, African Americans taught or socialized not to express emotion to people outside of their race may have difficulty sharing their feelings or seeking treatment from a European American.

Essentially every construct measured by psychological assessments is influenced by culture. Unfortunately, there seems to be a general assumption that constructs manifest identically and therefore can be measured identically for people of all cultural backgrounds. There is a Eurocentric bias in assessment research, which has impacted test construction and content (Dana 2000). Many aspects of the assessment process are based on Eurocentric cultural beliefs and values, which can put African Americans, particularly those who are more aligned with Africentric cultural values, at a disadvantage (Helms 1992). Individualism is a value that is apparent in assessment. Eurocentric culture values individual achievement and holds individuals responsible for their own successes and failures (Helms 1992). Interpretations of assessments from such a perspective may consider the person to be defective in some way, discounting social factors and disadvantages that may impact functioning. Further, Africentric beliefs and values focus on communalism, and the idea that one's assessment scores are representative of African Americans as a whole may influence assessment performance (Helms 1992).

Considering the consequences of both an accurate and inaccurate assessment, it is imperative for clinicians to understand and incorporate cultural contexts and nuances into assessment instrumentation and conclusions made from assessments. The rest of this section will focus on a brief overview of pertinent cultural issues that may influence the assessment process when working with African Americans.

Ethnicity Differences To date, assessment research that investigates the heterogeneity of those labeled as Black or African American is limited. Black people of different ethnicities or immigration statuses including Caribbean Americans, recent immigrants from Africa, second or third generation immigrants, and descendants of African slaves have often been treated as if they are one homogeneous group even though there is some indication that these groups may have differential results with psychological assessments. Specifically, Caribbean born Americans and US born Americans have been found to perform differently on cognitive performance tests (Byrd et al. 2005a; Gonzalez et al. 2007). Different worldviews and systems of education could contribute to these findings but it is clear that an individual's ethnic heritage is an important issue to consider and research when assessing people of African descent.

Education and Socioeconomic Status Education attainment and socioeconomic status are highly related and are essential considerations when assessing African Americans. Of African Americans that are 25 and older, about 82 % have a high school diploma and about 18 % have a bachelor's degree. Additionally, a little over one quarter of African Americans are at or below the poverty level (U.S. Census Bureau 2012). There has been a focus on correcting assessment norms based on completed years of education but this approach has been found to overestimate the quality of African American educational experiences by assuming that getting a high school diploma at a well funded school with a small student-teacher ratio is the same as getting a diploma at a grossly under-funded and overpopulated school. Recent studies suggest that reading skill may be a better indicator of educational experience (Byrd et al. 2005b; Manly et al. 2002; Ryan et al. 2005) and reading level has been found to attenuate racial differences between African Americans and European Americans on neuropsychological tasks (Manly et al. 2002; Rohit et al. 2007; Ryan et al. 2005).

Familial Structure The African American family cannot be simply defined but it is clear that trends differ for African Americans compared to other ethnic groups. According to Census data, 34 % of African Americans are married and 22 % are widowed, divorced, or separated. Additionally, approximately 40 % of African American men and women have never been married, which is the highest proportion of any racial category (U.S. Census Bureau 2003). Nearly one half of African American families have a female leader only and 75 % of all African American children in the last two decades are likely to live for some portion of their childhood with only mothers (Bumpass and Sweet 1989). These characteristics may make some think that family is unimportant in this group (Moynihan 1965) but the opposite seems to be true. Nobles' (1985) model of the African American family suggests that family is characterized by continual flexibility in the extended kinship, role flexibility, strong mother roles and family survival, emphasis on the children/motherhood, communalistic socialization of children, spiritual over material values, respect for the elderly, and humanness. In clinical assessment, it is quite possible for the closeness and communalistic nature of many African American families to be considered maladaptive, dependent, or enmeshed.

Cultural Orientation Cultural orientation subsumes constructs such as racial identity, ethnic identity, and acculturation. Particularly, African Americans can vary greatly as a function of how much they identify with their race, how important their ethnicity is to them, and how much they have adopted the culture of the dominant society. Though research investigating the role of cultural orientation in assessment is in its infancy, there has been some indication that these constructs could influence assessment performance. Acculturation in African Americans has been found to influence neuropsychological assessment performance. Specifically those who endorse values and behaviors traditionally linked to African Americans have lower scores on measures of verbal skills (Manly et al. 1998), executive functioning (Kennepohl et al. 2004), psychomotor speed (Kennepohl et al. 2004), and memory (Manly et al. 1998). These findings may suggest that those who have adopted Eurocentric values and cognitive styles have an advantage on assessments. These results are not surprising in light of the fact that the development of the most commonly used psychological tests and clinical assessment instruments have often come from a Eurocentric worldview and standardized within European American populations.

Religiosity and Spirituality Both formalized religion and a general belief in a higher power have had a long standing role in the lives of many African Americans. African American culture is known to place a strong emphasis on religious beliefs and activities and strong religiosity has been viewed as an expression of African American cultural identity (Koenig et al. 2001; Levin et al. 1995). Spiritual beliefs or religious behaviors have the possibility of being misinterpreted and sometimes pathologized by clinicians who do not adequately assess how these behaviors and beliefs may influence thoughts about mental health, emotional expression, or cognitive ability.

Recommendations for Assessing African Americans

The following recommendations are presented as strategies for enabling greater cultural sensitivity and competency when preparing, selecting, administering, and interpreting psychological assessments for African Americans.

Preparation Phase The first recommendation may be the most important as well as the most difficult: Know thyself! It takes considerable self-reflection and awareness to grasp how culture influences day to day lived experiences. Thus, the first step in understanding others as cultural beings is to uncover our own cultural norms, attitudes, and traditions. Cultural self-assessment questions when working with African Americans could include: (1) What is my perception of African Americans?, (2) What steps do I need to take in order to articulate these perceptions to develop more awareness?, and (3) What kinds of information or resources do I need to acquire to increase my effectiveness with carrying out this assessment? (Bromley 1998).

When preparing the full assessment battery, clinicians should research and incorporate short, clinician friendly questionnaires (with suggested follow up qualitative questions) that could be used to assess variation in sociocultural experiences. Specifically, these questionnaires should not only include basic demographic information but they should be useful in gathering qualitative information that can influence interpretation of the assessment. Important areas to assess include the client's educational environment, cultural orientation, familial history, ethnicity, experiences with racism/discrimination, cultural mistrust, and the role of religion/spirituality in an individual's life. Using interviews and culturally oriented questionnaires can also enrich an assessment report by offering indicators of the client's strengths or environmental mastery even if it does not directly measure what the standardized psychological measure is supposed to measure. This way, factors that promote African American resilience and mental health could also be a component of the assessment procedures. Also, additional measurement of reading level should be assessed to aid in choosing the appropriate assessment instrument or as a lens for interpreting the results.

Assessment Selection First, assessment instrumentation should be chosen wisely and purposefully. If possible, use multiple methods of assessment when constructing assessment batteries for African Americans. Clinicians who use just one method of assessment are more likely to draw inaccurate conclusions (Meyer et al. 2001). Therefore, a combination of qualitative and quantitative methods that includes the use of a clinical interview with questions concerning the influence of cultural factors, behavioral observations, and/or one or more standardized or nonstandardized testing instruments is ideal (Ridley et al. 2001). Areas of psychological functioning should be assessed with various instruments and more data points may be needed to make more definitive inclusions. For example, clinicians could make it a regular practice to get observer data or to request permission to interview important people in the client's life. If standardized instruments are used, those that have been standardized with samples that included a substantial number of African Americans are preferred when available. The use of nonverbal intelligence measures may be preferred when the client's dialect is of concern. For clients of African descent for whom English is not their first language, the use of measures that have been translated and back translated, when available, can be used. In some cases, it may be beneficial to use a professionally trained interpreter (Padilla 2001).

Assessment Administration It is unlikely that the client-provider power differential will become nonexistent, so it is imperative that providers recognize it and take steps to prevent abuse of that power. Similarly, level of familiarity with a client is not something that can be changed within the context of an assessment session. However, building rapport and trust through a culturally competent interview can likely be helpful in alleviating some of the effects associated with lack of familiarity and cultural mistrust. Examiners can help to attenuate the effects of stereotype threat by asking questions during the interview that get at multiple social identities for African Americans, as opposed to priming with questions that are solely associated with race and ethnicity (Ambady et al. 2004). To deal with issues related to language, it is important to admit when there is uncertainty and ask for clarification. It is also helpful to check in with clients to ensure understanding. Creating an environment of open communication and asking clients if they have questions regarding the assessment can be effective mechanisms for building trust.

Nonverbal communication should be a central component of assessment with African Americans. Clinicians should be sure to document nonverbal behavior and signs of apprehension during the evaluation process because these behaviors could be related to mistrust or culturally insensitive questions or stimuli. By documenting nonverbal behaviors and inquiring further about these behaviors after the evaluation phase has been completed, examiners may enrich their information about the client and begin to form an appropriate lens to later interpret the assessment results. Additionally, there may be regional nuances that could only be uncovered with systematic recording of individuals' reactions or interpretations of items.

Assessment Interpretation Clinicians should examine research on the assessment measures used to determine if there is any evidence of selection bias with regard to African Americans and interpret assessment results accordingly. When a clinician has reached the point of consolidating the assessment information and results, the report should be as comprehensive as possible. Assessment reports should include culturally-relevant information about clients, including ethnic identity, religiosity, acculturation, and cultural mistrust. Also, given the culturally-relevant information gathered about an individual, the examiners should include information about how the given test results could be related either directly or indirectly to cultural factors, personal strengths that the individual may have, as well as provide culturally relevant recommendations.

Conclusion

Although cross-cultural assessment of psychological phenomena is not a new endeavor and although the numbers of African Americans have been growing steadily, our ability to assess them in culturally appropriate ways has not kept pace. Instruments for the major constructs in psychology—e.g., personality, intelligence, psychopathology—tend to still be limited to those that have been developed on primarily European American samples. An overemphasis on internal validity to the virtual exclusion of external validity in the development of assessment instruments has led researchers to assume the generality of their constructs and instruments rather than demonstrating it. Researchers tend to deemphasize background characteristics and assume their assessment instruments and the theories that underlie them are universally applicable. We fail to follow good scientific principles when we assume that findings from research on one population can be generalized to other populations; generality should be empirically tested and considerable research shows that when generality is tested, important ethnic and cultural differences are often found (Sue 1999).

The considerable national, linguistic, ethnic, racial, cultural, and social diversity among African Americans has important implications for psychological assessment in several ways. First, it reflects that African Americans are a heterogeneous group, and not all of the people that get assigned that label identify with the label. Second, the generalizability of the tests we use must be demonstrated rather than assumed. A hallmark of our scientific method is scientific skepticism, the view that one must question, doubt, and suspend judgment until sufficient information is available (Sue 1999). As scientists, psychologists base their decisions on evidence rather than on prejudice, bias, or uncritical thinking. A third, and related, implication is that the tests we use to assess African Americans must be valid for them. Remember that tests do not possess validity. Rather, the validity of a test is a contextual factor that depends on how and on whom the test is used. In other words, validity refers to the soundness and defensibility of the interpretations, inferences, and uses of test results. It is the interpretations and uses that are either valid or invalid; thus, a test can be valid for one purpose but invalid for another. Consequently, we cannot assume that a test developed for one group will be valid when used with another group. Indeed, no test is valid for all purposes or in all situations. People who work with African Americans must ensure they are using tests that are appropriate for them based on the sample on which the instruments were standardized.

References

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology, 40,* 401–408.

American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bracken, B. A., & McCallum, R. S. (2001). Assessing intelligence in a population that speaks more than two hundred languages. In A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 405–431). San Francisco: Jossey-Bass.

Bromley, K. D. (1998). *Language art: Exploring connections*. Needham Heights: Allyn & Bacon.

Butcher, J. N., & Garcia, R. E. (1978). Cross-national application of psychological tests. *Personnel and Guidance Journal, 56*(8), 472–475.

Bumpass, L., & Sweet, J. (1989). National estimates of cohabitation. *Demography, 26,* 615–625.

Byrd, D. A., Sanchez, D., & Manly, J. J. (2005a). Neuropsychological test performance among Caribbean-born and U.S. born African American elderly: The role of age, education, and reading level. *Journal of Clinical and Experimental Neuropsychology, 27,* 1056–1069.

Byrd, D. A., Jacobs, D. M., Hilton, H. J., Stern, Y., & Manly, J. J. (2005b). Sources of error on visuoperceptual tasks: Role of education, literacy, and search strategy. *Brain and Cognition, 58,* 251–257.

Dana, R. H. (1996). Culturally competent assessment practice in the United States. *Journal of Personality Assessment, 66*(3), 472–487.

Dana, R. H. (2000). Culture and methodology in personality assessment. In I. Cuellar & F. A. Paniagua (Eds.), *Handbook of multicultural mental health* (pp. 97–120). San Diego: Academic.

Faris, H. (2013). http://www.americanprogress.org/issues/immigration/news/2012/12/20/48571/5-fast-facts-about-black-immigrants-in-the-united-states/. Accessed 17 Dec 2012.

Franklin, V. P. (2007). The tests are for the dogs: The Journal of Negro Education, African American children, and the intelligence testing movement in historical perspective. *Journal of Negro Education, 76*(3), 216–229.

Gonzalez, H. M., Whitfield, K. E., West, B. T., Williams, D. R., Lichtenberg, P. A., & Jackson, J. S. (2007). Modified-symbol digit modalities test for African Americans, Caribbean Black Americans, and non-Latino Whites: Nationally representative normative data from the National Survey of American Life. *Archives of Clinical Neuropsychology, 22,* 605–613.

Gopaul-McNicol, S., Reid, G., & Wisdom, C. (1998). The psychoeducational assessment of Ebonics speakers: Issues and challenges. *Journal of Negro Education, 67*(1), 16–24.

Guerra, N. G., & Jagers, R. (1998). The importance of culture in the assessment of children and youth. In V. C. McLoyd & L. Steinberg (Eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 123–142). Mahwah: Lawrence Erlbaum.

Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47*(9), 1083–1101.

Jones, J. M., & Rhee, E. (2004). The dialectics of race: Academic perils and promises. In J. M. Darley, M. P. Zanna, & H. L. Roediger III (Eds.), *The compleat academic: A career guide* (2nd ed., pp. 295–310). Washington, DC: American Psychological Association.

Kaplan, R. M., & Saccuzzo, D. P. (2001). *Psychological testing: Principles, applications, and issues* (5th ed.). Belmont: Wadsworth/Thomson Learning.

Kennepohl, S., Shore, D., Nabors, N., & Hanks, R. (2004). African American acculturation and neuropsychological test performance following traumatic brain injury. *Journal of the International Neuropsychological Society, 10,* 566–577.

Kent, M. M. (2007). Immigration and America's black population. *Population Bulletin, 62*(4), 1–16.

Koenig, H., McCullough, M., & Larson, D. (2001). *Handbook of religion and health*. New York: Oxford University Press.

Levin, J., Chatters, L., & Taylor, R. (1995). Religious effects on health status and life satisfaction among Black Americans. *Journal of Gerontology: Social Sciences, 50B,* S154–S163.

Manly, J. J., Miller, S. W., Heaton, R. K., Byrd, D., Reilly, J., Velasquez, R. J., et al. (1998). The effect of African American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *Journal of the International Neuropsychological Society, 4,* 291–302.

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society, 8,* 341–348.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., & Diess, R. R., et al. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*(2), 128–165. doi:10.1037//0003-066X.56.2.128.

Moynihan, D. P. (1965). *The Negro family: The case for national action*. Washington, DC: Office of Policy Planning Research, U. S. Department of Labour.

Neighbors, H. W., Trierweiler, S. J., Ford, B. C., & Muroff, J. R. (2003). Racial differences in DSM diagnosis using a semi-structured instrument: The importance of clinical judgment in the diagnosis of African Americans. *Journal of Health and Social Behavior, 43,* 237–256.

Nobles, W. (1985). *Africanity and the Black family: The development of a theoretical model* (2nd ed.) Oakland: Institute for the Advanced Study of Black Family and Culture.

Padilla, A. M. (2001). Issues in culturally appropriate assessment. In L. A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 5–28). San Francisco: Jossey-Bass.

Paniagua, F. A. (2005). *Assessing and treating culturally diverse clients: A practical guide* (3rd ed.). Thousand Oaks: Sage.

Ponterotto, J. G., Gretchen, D., & Chauhan, R. V. (2001). Cultural identity and multicultural assessment: Quantitative and qualitative tools for the clinician. In L. A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 67–99). San Francisco: Jossey-Bass.

Ridley, C. R., Hill, C. L., & Wiese, D. L. (2001). Ethics in multicultural assessment: A model of reasoned application. In L. A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 29–46). San Francisco: Jossey-Bass.

Rohit, M., Levine, A., Hinkin, C., Shogik, A., Saxton, E., Valdes-Sueiras, M., & Singer, E. (2007). Education correction using years in school or reading grade level equivalent? Comparing the accuracy of two methods in diagnosing HIV-associated neurocognitive impairment. *Journal of the International Neuropsychological Society, 13,* 462–470.

Ryan, E. L., Baird, R., Rivera-Mindt, M., Byrd, D., Monzones, J., Morgello, S., & The Manhattan HIV Brain Bank. (2005). Neuropsychological impairment in racial/ethnic minorities with HIV infection and low literacy levels: Effects of education and reading level in participant characterization. *Journal of the International Neuropsychological Society, 11,* 889–898.

Skiba, R. J., Knesting, K., & Bush, L. D. (2002). Culturally competent assessment: More than nonbiased tests. *Journal of Child and Family Studies, 11*(1), 61–78.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811.

Sue, S. (1999). Science, ethnicity, and bias: Where have we gone wrong? *American Psychologist, 54*(12), 1070–1077.

Terrell, F., Terrell, S. L., & Nickerson, K. J. (1996). Cultural identification and cultural mistrust: Some findings and implications. In R. L. Jones (Ed.), *Advances in Black psychology*. Berkeley: Cobb & Henry.

Terrell, F., Taylor, J., Menzise, J., & Barrett, R. K. (2009). Cultural mistrust: A core component of African American consciousness. In H. Neville, B. Tynes, & S. Utsey (Eds.), *Handbook of African American psychology*. (pp. 299–309). Thousand Oaks: Sage.

Terrell, F., Terrell, S. L., & Taylor, J. (1981). Effects of race of examiner and cultural mistrust on the WAIS performance of Black students. *Journal of Consulting and Clinical Psychology, 49*(5), 750–751.

U.S. Census Bureau. (April 2003)*The Black population in the United States: March 2002*. In J. McKinnon, Current population reports. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (4 Jan 2012) Facts for features: Black (African American) History Month, February 2012. U.S. Census Bureau News. http://www.census.gov/newsroom/releases/archives/facts_for_features_special_editions/cb12-ff01.html. Accessed 15 Oct 2012.

Whatley, P. R., Allen, J., & Dana, R. H. (2003). Racial identity and the MMPI in African American male college students. *Cultural Diversity and Ethnic Minority Psychology, 9*(4), 345–353. doi:10.1037/1099-9809.9.4.345.

3

Structured and Semistructured Clinical Interviews Available for Use Among African American Clients: Cultural Considerations in the Diagnostic Interview Process

Davor Zink, Bern Lee and Daniel Allen

Introduction

Even though the USA keeps growing more diverse every day, there is still a deficit in completely understanding how ethnicity and culture influence psychological assessment and more specifically, the interview process as well as diagnostic and treatment decisions (Aklin and Turner 2006). Symptom patterns associated with mental disorders do not manifest identically across cultures or people of different ethnicities. For example, research suggests that African American (AA) and Hispanic patients tend to exhibit more somatic and physical complaints when diagnosed with depression compared to Caucasians (Brown et al. 1996; Myers et al. 2002). In addition, different kinds of hallucinations are manifested in people of different cultures (Bauer et al. 2011). Further, it is plausible that the phenomenology of disorders differs across racial/ethnic groups, resulting in some cultures viewing behaviors as pathological that others regard as normative (Lewis-Fernández et al. 2010). For example, some cultures that are far removed from the Western culture consider hallucinations and epileptic seizures as mystical gifts or higher powers of a religious nature.

The Diagnostic and Statistical Manual of Mental Disorders Fourth Edition Revised (DSM-IV-R) acknowledged cultural differences in diagnosis by including an outline for preparation of cultural formulations, increasing the description of the client's individuality by multidimensional data collection and amalgamation of a clinical presentation format, and including a glossary of culturally bound syndromes (American Psychiatric Association 1994; Dana 2008). For example, among AAs "spell" refers to a trance state in which individuals can communicate with deceased relatives and which is associated with brief periods of personality change (American Psychiatric Association 1994). The Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5) also included a glossary of cultural concepts of distress. However, the glossary does not include cultural concepts of distress related to AAs. In addition, it updated the cultural formulation outline and added a Cultural Formulation Interview (CFI; American Psychiatric Association 2013).

These updates suggest that cultural issues now have more relevance for clinical diagnosis; moreover there is substantial research reporting the importance and influence of cultural and ethnic factors in psychological assessment, particularly in the clinical interview (e.g., Adebimpe 1981; Neighbors 1989; Paniagua 2001; Trierweiler et al. 2000; Aklin and Turner 2006; Grieger 2008). The principal goal of the clinical interview is to gather information related to the problem at

---

D. Allen (✉) · D. Zink · B. Lee
Department of Psychology, University of Nevada,
Las Vegas, 4505 Maryland Parkway, Las Vegas,
NV 89154-5030, USA
e-mail: daniel.allen@unlv.edu

D. Zink
e-mail: davor13nzp@gmail.com

B. Lee
e-mail: bernglee@gmail.com

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_3, © Springer Science+Business Media New York 2015

hand. It represents an interpersonal interaction that aims to elucidate an individual's feelings and attitudes about their symptoms and problem behaviors, while at the same time allowing the clinician to observe verbal and nonverbal behavior. It is the best tool to gather information that that leads to a diagnosis and establishes the stage needed for a fruitful therapeutic relationship (rapport) (Aklin and Turner 2006; Turner et al. 2003). When working with AAs it is also the best tool to gather and incorporate relevant cultural and ethnic factors that might influence the assessment process, case conceptualization, and diagnosis.

Types of Interviews

Clinical interviews remain one of the most commonly used methods to assist in the diagnoses of an individual. There are three main types of clinical/diagnostic interviews: open, structured, and semistructured.

Open Interviews

Open interviews permit clinicians to choose what questions to ask, the depth to which topics are explored, the length, or any other pertinent variable. They are advantageous in that the interview can be tailored to the specific concerns of the patient, they can be less time consuming, are generally easy to learn and administer, can be administered anywhere, the conversational nature aids rapport, and they allow for flexibility of depth of symptom exploration. Open interviews risk intrusion of theoretical or personal biases, omission of important questions, variation of wording and question tone, variability of the order in which the questions are asked, variation of the depth and style of recording, and often lack ratings to measure presence and severity of symptoms (Rogers 2001). These variables can affect how patients respond to diagnostic questions. In a seminal paper Ward et al. (1962) reached the conclusion that the majority of diagnostic variability is a result of the evaluations, not the patients. In his study, 62.5 % of diagnostic variability resulted from criterion variance, 32.5 % from information variance, and a scant 5.0 % from patient variance. Criterion variance refers to differences among clinicians in implementing standards for what is clinically important and when the diagnostic criteria are met. Information variance refers to differences among clinicians in the questions that are used, the observations made, and how the gathered information is organized. Patient variance refers to differences within the same patient that lead to significant discrepancies in clinical presentation and subsequent diagnosis. These sources of variance are likely to be present in open interviews. Adding variables such as culture, ethnicity or language barriers to the equation greatly risks further increased variability (Aklin and Turner 2006). In a study by Strawoski et al. (1997), when a patient was of minority status, information variance was the cause of diagnostic disagreements in 58 % of the cases. Criterion variance was the cause of disagreement in 42 % of the cases and was not associated with race.

Structured Interviews

Structured interviews allow for a systematic evaluation by standardizing the specific language of clinical questions, the order of these questions, and the quantification of responses. All questions must be asked verbatim as it is instructed. Standard questions and optional probes are usually utilized (Rogers 2001). Structured interviews may be advantageous because they reduce patient variance and capitalize on systematic evaluation to reduce misdiagnosis. Their comprehensiveness is likely to abate missed diagnosis by removing a priori hypotheses from the interview process, and a high level of standardization is ensured. In addition, they allow for systematic comparisons to be made (across collateral sources, time, clinicians, within the same patient, within settings, and within diagnosis; Rogers 2001). These qualities make structured interviews the instrument of choice for research purposes. Such instruments, however, are lengthy, require advanced training, and their rigid guidelines cannot account for all possible eventualities. Structured interviews enhance information variance, sometimes at the cost of criterion variance (Rogers 2001).

Semistructured Interviews

Semistructured interviews include guidelines that allow clinicians considerable flexibility while still assuring that a certain set of questions will be reviewed across clinicians and patients (Aklin and Turner 2006). They utilize standard questions, optional probes, and unstructured questions. Semistructured interviews tend to increase criterion variance, sometimes at the cost of information variance (Rogers 2001). Aklin to structured interviews, semistructured interviews have the advantage of standardization and reduced variability. In addition, examiners may ask their own questions when diagnostic issues remain unresolved. To improve patient understanding, they allow for a more spontaneous and conversational interview, and they can be modified according to a client's specific needs (Aklin and Turner 2006; Rogers 2001). Disadvantages include long administration time, extensive training requirements, and reliance on protocols. Research shows that structured and semistructured interviews have better validity and reliability when assessing individuals that belong to an ethnic minority when compared to open interviews (Widiger 1997). Structured and semistructured interviews allow for systematic and comprehensive coverage of symptoms by ensuring that specific criteria are met for diagnoses, and decrease variance resulting from differences in ethnicity and culture (Turner et al. 2003).

Cultural Considerations when Interviewing AAs

The following are important cultural factors specific to AAs that are important to consider in the diagnostic interview process, regardless of the type of interview utilized.

Socioeconomic Status and Population Characteristics

Currently AAs represent 13 % (38.9 million) of the total US population. Their poverty rate (13 %) is close to being two times higher than that of all households (25.5 %) and their unemployment rate (13.6 %) is almost twice as high (7.4 %) (U.S. Census Bureau 2010). Nearly 13 % of AA adults do not have a high school diploma (Fry 2010). In addition, the median wealth of white households is 20 times that of AA households and nearly half of all prisoners in the USA are AAs. Prison inmates and people experiencing poverty are at high risk of developing a mental illness (NAMI 2004). These negative statistics are influenced by the disadvantaged position of AAs, racism, and poverty. However, these numbers do not represent the whole of the AA population in the USA. Most of the available literature is based on the economically disadvantaged portion of the AA population; therefore the diversity of this population is somewhat underreported (Holmes and Morin 2006; Sue 2013). For example, over 38 % of AA households are middle class vs. 44 % of all households (U.S. Census Bureau 2010). All of these facts and statistics should be taken into account when conducting a diagnostic interview with AAs; nevertheless they should not influence or bias the clinician's judgment. Socioeconomic status (SES) is a particularly important variable to consider when assessing minorities. For example, many clinicians tend to perceive individuals of a lower SES as having more problems than individuals of higher SES (Bentacourt and Lopez 1993; Lindsey and Paul 1989; Robins and Regier 1991; Snowden and Cheung 1990). Using unstructured interviews increases the probability that these kinds of beliefs will negatively influence the clinician's judgment and symptom interpretation (Garb 1997). Using structured and semistructured interviews could reduce this probability (Aklin and Turner 2003).

Racial Identity

Research suggests that members of cultural and ethnic minority groups go through a sequential process of racial identity (Sue 2013). For many AAs this process consists of an evolution from a non-Afrocentric identity to one that is Afrocentric. Cross (1995) presented a model of AA racial identity that involves the following stages: pre-encounter, encounter, immersion–

emersion, and internalization. These stages are associated with differences in views related to the self and relationships with others, beginning with the acceptance of white culture and rejection or devaluation of black culture and ending with an appreciation and acceptance of both black culture and aspects of white culture (Sue 2013). AAs that are at the pre-encounter level are less prone to report racial discrimination, while those in the immersion stage are usually younger and least satisfied with societal conditions (Hyers 2001). In addition, AAs at the first stage tend to prefer a white counselor, whereas those in later stages prefer an AA counselor (Parham and Helms 1981). Discussing racial identity during the interview process and getting a picture of which stage the patient is in could help build the therapeutic relationship and reduce the possibility of misinterpreting answers to certain questions.

Racism, Discrimination, and Stereotypes

Clinicians, like every other person, could potentially be racist towards a minority group or hold stereotypes or prejudices regarding a particular minority group. Diagnostic errors could be the result of direct racism or indirect racism. Indirect racism refers to preconceived notions about a certain group (Aklin and Turner 2006). Further, clinicians should actively try to avoid engaging in microagressions. "Microaggressions are brief and commonplace daily verbal or behavioral indignities, whether intentional or unintentional, that communicate hostile, derogatory, or negative racial slights and insults that potentially have a harmful or unpleasant psychological impact on the target person or group" (Sue et al. 2007). Even counselors with a significant background in cross-cultural practices may engage in micro-aggressions against AA clients (Gushue 2004), and these micro-aggressions significantly impact the therapeutic working alliance (Constantine 2007).

Regarding discrimination, Approximately 50 % of whites believe blacks have equal societal opportunities and 81 % of blacks believe more change is necessary (Pew Research Center 2010). Moreover, many AAs believe that racial profiling occurs frequently (Carlson 2004) and 43 % report there is a great deal of anti-black discrimination in the USA, compared to 13 % of whites. This lack of trust and feelings of discrimination towards the mainstream culture can affect perceptions of social and health systems among AAs (Miller et al. 2001). Further, generalizations about a certain ethnic or cultural group and perceptions influenced by stereotypes can negatively affect the diagnostic process (Whaley 1998). Common stereotypes related to AAs include, but are not limited to: they are unmotivated for treatment, they are violent, hostile, and they are inferior or possess less intelligence (Devine et al. 1991; Monteith et al. 1996). Classic social psychology research regarding stereotypes, prejudice, and discrimination becomes relevant regarding the diagnostic interview of a minority group member. According to Nisbet and Wilson (1977) people are often unaware of the range of factors that can significantly influence their judgments. Intergroup bias seems to be an important factor in this kind of subtle influence. Stereotypes and prejudice are activated so rapidly and spontaneously that they can color initial reactions and potentially bias the processing of subsequently encountered evidence. This could have a negative impact on the outcome of a diagnostic interview, regardless of the type, if the clinician is unaware of the influence of his own stereotypes on his judgments. In addition to influencing people's judgments, stereotypes can also influence people's behavior. When stereotypes become activated they tend to elicit a corresponding behavioral response (automatic behavior; Dijksterhuis and Bargh 2001). For example, if a clinician consistently engages in less eye contact and hand movements with AA patients compared to whites, this could influence the AA patients responses to diagnostic questions and their level of engagement without the clinician realizing it. Another manifestation of stereotypes and prejudice at the behavioral level is the self-fulfilling prophecy (Darley and Fazio 1980). In an interracial interaction, this occurs when the biased expectations (driven by stereotypes) people have regarding the other person lead them to behave in a way that will provoke the expected kind of behavior. In a clinical interview a self-fulfilling prophecy could lead the interviewer to a certain diagnoses despite of evidence suggesting that it is false.

Minority group members also hold stereotypes and prejudices and this can include beliefs about their own group (internalized racism). Stereotype threat refers to apprehension experienced by members of a group that their behavior might confirm a stereotype. In classic study Steele and Aronson (1995) showed that when AAs were told that they were going to take an intelligence test they performed worse compared to whites. In contrast, when they were not told about the nature of the test no differences in performance were found. Mentioning that the test was diagnostic of intelligence triggered the stereotype that AAs are academically inferior to whites, which led to poorer performance. Clinicians have to be careful not to elicit this kind of reaction in their AA patients with the questions they ask. When assessing patients, vague and complicated symptoms and complaints, in addition to the clinician's poor cultural competence enhances stereotypes in clinical settings. Because of this, mental health professionals tend to inappropriately link stereotypical beliefs and thoughts to AAs with mental illness characteristics, behaviors, and factors (Neighbors et al. 2003). Using structured and semistructured interviews diminishes the clinicians' opportunities to include generalizations about a particular person or group and the influence of stereotypes when making diagnostic decisions (Aklin and Turner 2003). According to Devine (1989) automatic stereotypical reactions are likely to prevail unless they are suppressed by controlled process. Mental health professionals have the responsibility to evaluate themselves and realize what kind of stereotypes, prejudices, and racist beliefs they hold before attempting to serve a minority group such as AAs.

Language

It has been argued that the diagnostic bias inherent in unstructured interviews are to a large extent mediated by a clinician's freedom to fall prey to stereotypes and prejudice. These biases are to some extent reined in by a more regimented style in structured and semistructured interviews and greater cognitive focus may be dedicated to objective evidence of diagnostic relevance. Clinicians are trained to pay close attention to language. This makes intrinsic sense, as linguistic fluidity, direction of speech, content of speech, poverty of speech, fluidity, and so forth are all significant cues to making diagnostic distinctions, many of which have great impact on the lives of a client. Given that language and culture are heavily intertwined, the importance of understanding the linguistic aspects of the client's culture, and indeed unique subculture, cannot be overemphasized. Language is an important cultural factor that influences the outcome of a diagnostic interview. Language barriers and dialects used could affect how a patient answers diagnostic questions. For example, if a clinician is not familiar with common slang terms used by AAs, something that the patient says could be misinterpreted. Language capabilities, use, and preference of the patient are also important variables to consider. Research suggests that these variables largely depend on SES (Al-issa 1995). Structured and semistructured interviews seem to reduce communication errors because they provide the clinician with prompts that have to always be read the same way and they are commonly dichotomously coded, leaving the clinician less room to misinterpret verbal statements (Miller et al. 2001). Nevertheless, before using a structured or semistructured interview it should be considered whether the language that the instrument was developed with is appropriate for the particular patient being assessed.

Diagnostic Bias

Cultural bias in diagnosing psychiatric disorders, when assessing minorities, have been often reported in the literature (Adebimpe 1981; Bell and Mehta 1980; Neighbors et al. 1999; 2003; Strakowski et al. 2003). A common criticism is that current diagnostic systems (e.g., DSM-5) were developed primarily for individuals with a Western cultural background; therefore they fail to adequately include cultural differences that could impact diagnoses. Valid and reliable diagnoses are important for several reasons, including, but not limited to establishing prevalence rates, prescribing appropriate treatment, and identifying individuals at risk for developing psychiatric disorders. In addition, there is considerable stigma related to a psychiatric diagnosis, this may cause ethnic minority groups to be reluctant to access mental health services when more severe diagnoses are made (McGuire and Miranda 2008).

Regarding diagnostic bias towards AA patients there are several studies in the literature that report that AAs are more likely to be diagnosed with schizophrenia (SZ) compared to whites (Barnes 2008; Bresnahan et al. 2007; Minsky et al. 2003; Trierweiler et al. 2000; Strakowski et al. 1996), with some studies suggesting AAs are up to three times more likely to be diagnosed with SZ compared to whites (Bresnahan et al. 2007; Eack et al. 2012). In a classic review, Adebimpe (1981) suggested the high rates of SZ diagnoses in AA were due to the following variables: cultural distance between patient and provider (e.g., differences in language, values, and expressions of distress), stereotypes of AA psychopathology (e.g., hostility, reluctance to get treatment), false-positive symptoms (e.g., flat affect, paranoia), and biased diagnostic instruments (not culturally sensitive). DeCoux Hampton (2007) expanded on these factors and suggested that several client-level, care process-level, and system-level variables likely contribute to diagnostic bias, with the manner in which care is accessed (system level) being particularly important. AAs tend to use emergency and acute care services more than other racial and ethnic groups; therefore greater symptom severity and acute psychotic episodes may increase the possibility of a diagnosis of SZ (DeCoux Hampton 2007).

Strakoswki (1996) examined the relative importance placed on different symptom clusters when diagnosing SZ. AAs in this study had higher attributions of auditory hallucinations. Authors suggested that clinicians tend to diagnose SZ in AAs based in criterion A symptoms. In line with these results, Strakoswki (2003) examined diagnostic patterns in interviewers who were blinded and unblinded to race. Participants were 195 inpatients with psychotic symptoms. The Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID) III was utilized for diagnostic purposes. Results showed unblinded interviewers endorsed criterion A or auditory hallucinations in AAs, while blinded clinicians found no ethnic differences. The symptoms identified by blinded and unblinded interviewers were similar, but unblinded clinicians reported higher rates of SZ among AAs (47 % compared to 17 % in whites). More recently, Eack et al. 2012 conducted a study that investigated whether the clinicians' perception of the client's honesty in reporting symptoms influenced diagnostic decisions. In this study the DSM-III-R Criteria Checklist (Janca and Helzer 1990) was used to determine psychiatric diagnosis. The DSM-III-R Criteria Checklist is a structured interview and checklist that obligates the clinicians to go through all Axis I diagnostic criteria before making a final diagnosis. The Brief Psychiatric Rating Scale (BPRS; Overall and Gorham 1962) was used to determine psychiatric symptomatology. In line with previous research AAs were diagnosed with SZ more than twice as frequently compared to Caucasians (45 vs. 19 %), in addition clinicians perceived AA clients to be less honest than whites. Results indicated that diagnostic biases were substantially reduced after controlling for perceived honesty. In line with these results, Trierweiler et al. (2000) examined clinicians' attributions associated with diagnosing SZ. They used an open interview to determine diagnosis, and results showed that hallucinations and paranoid suspicious attitudes were more often attributed to AA patients, whereas elevated mood and the combination of negative symptoms and dysphoric mood were more often attributed to non-AA patients. They concluded clinicians tend to use different criteria to diagnose SZ depending on the client's race (Trierweiler et al. 2006). Consistent with these results, Neighbors (2003) reported clinicians link symptom observations to diagnostic categories differently depending on the client's race. In this study the sample was 665 psychiatric inpatient AAs and whites. A shortened version of the DSM-III-R Criteria Checklist was used to determine psychiatric diagnoses. Results showed AAs were somewhat more likely to be diagnosed with SZ while whites were much more likely to be diagnosed with bipolar disorder. In addition, loose associations, inappropriate affect, auditory hallucinations, and vague speech were more predictive of a SZ diagnosis in AAs than Caucasians. Even though, the frequency of these symptoms was the same in both groups, the importance the clinicians attributed to them was different depending on race. This influenced whether a diagnosis of SZ or bipolar disorder was given (Neighbors 2003). In a similar study, Neighbors (1999) administered open and semistructured interviews to a sample of AAs and Caucasians psychiatric inpatients (291 patients for phase 1/hospital and 665 patients for phase 2/research). Findings indicated that when patients were evaluated in a hospital setting using clinician-structured interviews (open) AAs were more likely to be diagnosed with SZ and less likely to be diagnosed with a mood disorder compared to whites. When the DSM-III-R Criteria Checklist was used to determine diagnosis the disparity in diagnosis across groups was still present, however it was largely reduced. These results suggest that the use of a semistructured diagnostic instrument can help reduce clinicians' diagnostic bias due to cultural differences. Notably, in studies that used a semistructured interview to determine diagnosis (SCID III) to address clinicians' subjectivity, AAs were still more likely to be diagnosed with SZ than whites, even though the disparity was sometimes reduced (Strakowski et al. 2003; Trierweiler et al. 2000). In contrast with these results, Jeste et al. (1996) examined racial differences in the prevalence of SZ in a community sample of AAs and whites and found no differences based on the Brief Psychiatric Rating Scale (BPRS). Providing support for these results, in another study with 226 AAs, Latinos, and Whites with SZ, also using the BPRS for diagnoses, no racial differences were found in symptomatology (Bae and Breke 2002).

The influence of the clinician race in diagnosing AAs with SZ has also been studied. Matthews et al. (2002) conducted a study with a large sample of inpatients with psychotic and mood disorders that compared the diagnoses of culturally matched and unmatched patients and clinicians. AAs were diagnosed with SZ and schizoaffective disorder more often compared to whites, regardless of the clinicians' race. In another study with 234 inpatients, AA providers were significantly less likely to diagnose mood disorders in AA patients compared to non-AA providers (Trierweiler et al. 2005). In a similar study with 292 inpatients, AA clinicians were significantly more prone to diagnose SZ when hallucinations were present and non-AA clinicians were more likely to diagnose SZ when negative symptoms were present (Trierweiler et al. 2006).

In a recent brief review regarding diagnostic bias among AAs, Escobar (2012) suggested that the higher rates of SZ diagnosis among AAs in the USA may be due to clinicians overevaluating psychotic symptoms and minimizing affective disturbances when making diagnoses. Further, he suggested that in diagnosing SZ clinicians' bias may include factors of discrimination and stigma. White clinicians could perceive black patients with suspicion and fail to understand cultural nuances that could give hints about other diagnoses (Escobar 2012).

There is considerable research literature suggesting that AAs tend to be diagnosed with mood disorders less often than whites (e.g., Bell and Mehta 1980; Rollman et al. 2002; Breslauet al. 2006; Jimenez et al. 2010; Williams et al. 2007). Jonas et al. 2003 examined racial differences in depression diagnoses in a sample that included over 7000 adults. Results indicated AAs were more likely to be diagnosed with dysthymia, while whites were more likely to be diagnosed with major depressive disorder (MDD). The Diagnostic Interview Schedule (DIS) imbedded in a larger battery was used for diagnostic purposes in this study. In line with these findings, Woodward et al. (2013) reported older whites and Caribbean Blacks had significantly higher lifetime prevalence rate of MDD than AAs. However, no racial differences were found in 12-month prevalence rate of MDD. This study used data form a large epidemiological study that utilized the Composite

International Diagnostic Interview (CIDI) for diagnostic purposes. In contrast with these findings, studies conducted in primary care settings have found similar rates of depressive symptoms between AAs and whites (Brown et al. 1999; Diala et al. 2001; Oquendo et al. 2001). The difference in the rate of diagnosing depression between AAs and Caucasians could be related to several factors. One of them is the manifestation of depressive symptoms among AAs. AAs have been found to report more physical symptoms of depression instead of mood symptoms (Brown et al. 1996). Another factor is that there might be a higher prevalence of psychological protective resources among AAs, such as religion (Taylor et al. 2001; Chatters et al. 2008) and familial support (Chatters et al. 1985; Woodward et al. 2008). Further, research suggests that AA children are socialized to cultivate a high level of tolerance to unfair acts (e.g., racism, low SES, living in high-crime neighborhoods; Thornton et al. 1990). This type of upbringing could result in a protective psychological resource that buffers the manifestation of MDD.

Racial differences in the prevalence of posttraumatic stress disorder (PTSD) have also been reported, with minorities, particularly AAs and Hispanics having a higher prevalence of PTSD compared to whites (e.g., Lonigan et al. 1991; Norris 1992; Green et al. 1990; Kulka et al. 1990; Sutker et al. 1995). Graves et al. (2011) investigated the characteristics of PTSD in AAs in the primary care setting. They screened over 700 AAs for PTSD using the SCID IV and the Clinician Administered PTSD Scale for DSM-IV (CAPS) and concluded that most AA adult primary care patients in this sample were either undertreated or underdiagnosed. In a study that examined the influence of race in PTSD treatment, AAs were found to be significantly less likely to complete treatment compared to whites (Lester et al. 2010). In contrast with these findings Monnier et al. (2002) conducted a study that examined racial differences in outpatients seeking treatment for PTSD at a Veterans Affairs Medical Center. Participants were 71 Caucasians and 40 AAs and they were administered the CAPS in addition to an open clinical interview and other self-report measures for diagnostic purposes. Between racial groups no significant differences were found regarding anxiety, paranoia, dissociation, SZ, depression, and PTSD symptomatology. The researchers concluded that AAs and white veterans with combat-related PTSD did not differ in terms of severity and manifestation of symptoms. In a similar study using the CAPS that investigated symptom patterns and service use among AA and white veterans at a Veterans Affairs outpatient PTSD clinic, very few significant differences among groups were found. This suggests that white and AA veterans do not differ in the manifestation of PTSD and in the use of services (Frueh et al. 2004). In a critical review regarding racial differences in combat veterans with PTSD, Frueh (1998), suggested that the disparity in rates of PTSD between racial groups could be a function of distinct rates of traumatic stressors and other preexisting conditions. In addition, the general paucity and methodological limitations in the empirical data in this body of research considerably limits the conclusions that can be reached.

Cultural Mistrust

There is a documented trend of cultural mistrust among AAs (Nickerson et al. 1994; Terrell and Terrell 1984; Whaley 2001a). Whaley (1998) conducted a study that examined the experience of paranoia in AAs from a cross-cultural perspective. Participants included 96 depressed patients, 65 patients with SZ-like disorders, and 404 community members. Findings indicated AAs with and without a psychiatric diagnosis scored significantly higher on measures of distrust and perceived hostility by others. The author suggested that mild paranoia was a result of cultural mistrust. Cultural mistrust is defined as guardedness toward the dominant culture related to discrimination experienced by AAs. It stems from a long history of discrimination and racism towards AAs and the previously described long-held stereotypes and beliefs about this minority group. In relation with this cultural mistrust, AAs tend to underutilize mental health services compared to other ethnic and cultural populations (Whaley 2001b). AAs often might have negative feelings toward the mental health system and varying degrees of mistrust regarding Caucasians in the mental health system. Mistrust of mental health could influence current perceptions among AAs and subsequently affect the interviewing process (Aklin and Turner 2006). According to Whaley (1998) mild paranoia is higher in individuals of lower education and income and could be misinterpreted as psychopathy by clinicians. Further, misinterpretation of paranoid symptoms related to self-esteem protection plays a major role in misdiagnosing depressed individuals as schizophrenics. Among AAs cultural paranoia can be a nonpathological or normative behavior that functions as protection against racially based threats and discrimination (Ridley, 1984). Cultural mistrust being misinterpreted by clinicians as a symptom of psychosis instead of depression may contribute to the higher prevalence of SZ and lower prevalence of MDD diagnoses among AAs (Whaley 1998). This mistrust brings to light a significant assessment concern. Such lack of trust may mean that those who seek help are likely to be those in more acute distress. This makes it challenging to assess the actual base rates in a population of certain disorders thus affecting a clinician's judgment of population base rates. Further, clinicians may misinterpret such mistrust as paranoia or noncompliance, even when structured or semistructured interviews are employed for diagnosis. Clinicians should be aware that many AAs may experience apprehension when going to receive services and not let this phenomenon influence their diagnostic decisions. Clinicians that recognize a distrustful attitude in their patients should be flexible in their approach (Whaley 1997). Engaging in an honest conversation with the patients regarding their cultural values, experiences of racism, and mistrust could be helpful in building the therapeutic relationship and avoiding diagnostic errors (Aklin and Turner 2006).

Brief Review of Commonly Used Structured and Semistructured Interviews

Structured and semistructured clinical interviews traditionally demonstrate good reliability in the diagnosis of mental disorders. Such interviews, however, require a great deal of clinical acumen to operationalize properly, and therein lies one of their weaknesses for cross-cultural diagnosis; clinicians have a long, well-documented history of overdiagnosing severe psychopathology, and notably psychotic disorders in AA clients (Bell and Mehta 1980; Neighbors et al. 1999; 2003; Strakowski et al. 2003). The importance of proper use and interpretation of psychodiagnostic interviews cannot be overstated. The following are commonly used diagnostic structured and semistructured interviews for Axis I disorders available for clinicians to utilize with AAs.

Structured Clinical Interview for the DSM-IV (SCID)

The SCID (First et al. 2002) is considered by many as the gold standard for Axis I psychiatric diagnoses. This semistructured interview is designed for use with adults who have an eighth grade or higher reading level. There are several versions of the SCID, with versions developed for use in research (SCID-I-RV) and clinical settings (SCID-CV), although the shortened clinical version limits the diagnostic scope. The SCID-I-RV is broadest in scope and allows subtype specification, severity, and course specifiers. There is a patient edition (SCID-I/P) designed specifically for individuals with psychiatric disorders in research settings, as well as a nonpatient version (SCID-I-N/P) for use with participants who are not identified as having a psychiatric disorder. In addition, a draft version of the SCID has also been developed to cover childhood disorders (KID-SCID).

The SCID is a long, in-depth interview designed to screen for the great majority of DSM-IV diagnoses and strives to be unbiased in assessing diagnostic criteria. Administration time is approximately 1–2 h, however subsections of the SCID may be used in isolation when specific diagnostic concerns warrant precise investigation. This being said, full SCID versions may be inefficient as a broad-range diagnostic tool outside the research environment. A thorough SCID interview by an experienced clinician will still take well over 1 h for most clinical cases. As we transition into the SCID DSM-5 version, a more in-depth screening module is used to reduce total administration time. This being said, it is unlikely that the SCID will become practical as a whole for administration outside the research setting for most clients.

In its many forms, the SCID is perhaps the most widely used diagnostic research instrument in the English language, and rightly so. It offers an efficient, yet thorough treatment of diagnosis for a multitude of Axis I disorders, is modularized and thus can be tailored to fit the scope of the assessment, and is easy to learn and use for someone well versed in psychopathology. The SCID has been compared to clinically derived diagnoses (Shear et al. 2000) as a metric for validity of clinician diagnoses of nonpsychotic disorders in a large catchment study, and very low kappa coefficients (0.15) for overall agreement between SCID and charted diagnoses were attributed to the SCID's superiority as a diagnostic method. Similarly, Steiner et al. (1995) found poor reliability for the DSIII-R version of the SCID when compared to clinical diagnoses, also touting the merits of structured interviews. As with many structured interviews, the SCID has acceptable-to-excellent interrater reliability, with adequate administrator training (Lobbestael et al. 2011; Ventura et al. 1998).

Applications with AAs

Clinical judgment is a strong requisite when conducting semistructured interviews. The name "Structured Clinical Interview" belies the freedom a clinician has when making judgments. As such, one can be lulled into a false sense of security when making diagnostic distinctions using the SCID and fail to take into consideration specific and highly relevant cultural considerations that may prove important in the diagnosis, or rule-out, of any of a variety of mental disorders. Similarly, the structure of this diagnostic system is such that it cannot prevent the overinterpretation of the expression of psychotic symptomatology as a function of race; a phenomenon frequently observed in the diagnostic literature. Research on prevalence rates of eating pathology demonstrates significantly different attitudes toward body image and maladaptive eating patterns among AAs than among whites in the USA (Akan and Grilo 1995). Failing to consider the context in which a diagnostic criterion may or may not cause significant impairment in social or occupational functioning risks misdiagnosis.

Few studies report on validity or reliability measures of the SCID for AA clients. Those that do report a significant minority of AA clients also find (similarly to other studies of the SCID vs. other diagnostic strategies) abysmal concordance with clinical diagnoses given by open interviews (see Steiner et al. 1995). It is therefore challenging to tease apart the flaws in nonstructured clinical interviewing as opposed to those of a more structured interview format. The strong concordance between SCID diagnoses and those rendered by the MINI and CIDI (See Sheehan et al. 1998, 1997) ameliorate much concern with regard to the SCID's validity as a diagnostic instrument, though such lack of in vivo clinical concordance is grounds for pause.

The Diagnostic Interview Schedule (DIS)

The Diagnostic Interview Schedule version IV (DIS-IV; Robins et al. 1995) is a completely structured diagnostic interview developed to assess current and lifetime presence of mental disorders included in the DSM-IV. It is organized in 19 diagnostic modules that cover a wide variety of Axis I disorders. Diagnoses include substance use disorders, SZ, mood disorders, anxiety disorders, and a small selection of other disorders and those originating in childhood. The DIS can be administered by professionals and nonprofessionals with 1 week of training. Administration time is approximately 2 h. Questions are read verbatim and no opportunity is given for unstructured questions. A training manual that describes how to reliably code the clinical ratings of specific items is included (Robins et al. 1991). Items are scored in a format that combines clinical relevance and possible etiology. In addition, interviewers have to make ratings about the onset, duration, and recency of symptoms. Some of the DIS's fundamental characteristics are that it attempts to elucidate any organic etiology (exogenous substances, medical conditions), it includes the Mini Mental Status Examination (MMSE; Folstein et al. 1975), and it allows researchers to conduct comparisons across diagnostic systems through the inclusion of older DSM criteria, Feighner criteria, and Research Diagnostic Criteria (Rogers 2001).

The DIS was originally designed as a research tool for large epidemiological studies (Epidemiologic Catchment Area Program) to assess the prevalence and incidence of mental disorders in the USA (Regier et al. 1984). Several computerized versions of the DIS have been developed, although these have less diagnostic coverage. In addition, there is a shortened paper and pencil version that can be self-administered and covers depressive, anxiety, and alcohol-use disorders (Kovess and Fournier 1990). The DIS-IV has been translated into several languages, including Spanish and Chinese, which makes it a good tool for cross-cultural applications and research. The DIS has well-established validity and reliability for pre-DSM-IV diagnostic systems, making generalization to current diagnostic systems challenging. Overall, research shows moderate-to-good reliability and validity for the DIS (e.g., Helzer et al. 1985; Wells et al. 1988; Hesselbrock et al. 1982; North et al. 1997).

The DIS is limited in that it focuses more on etiology than symptom severity, it emphasizes diagnosis over symptom evaluation, and research suggests that it is vulnerable to response styles (Alterman et al. 1996, Cottler 1998; Rogers 2001). The DIS demonstrates utility in screening large samples of people for undetected mental disorders (epidemiological research), can be administered by nonprofessionals, and has been translated and validated in several languages. In addition, there is a children's version of the DIS: the Diagnostic Interview Schedule for Children (DISC; National Institute of Mental Health 1991; Columbia DISC Development Group 1999).

Applications with AAs

Robins et al. (1984) conducted a large epidemiological study using the DIS. Results from 9543 participants indicated that AAs had a higher prevalence of mental disorders than those of other races across all three sites. Few other diagnostic differences were found inconsistently across sites. In most DIS studies racial differences in diagnoses or the utility of the tool with AAs was either not explored or was not the main focus of the study, even when AAs comprised a large part of the sample. North et al. (1997) conducted a study comparing the diagnostic utility of the DIS and an open clinical interview in two mental health clinics for homeless people; samples were 75 and 69 % AA, respectively. Compared to the clinical interviews, diagnosis made with the DIS underdiagnosed antisocial personality disorder and overdiagnosed MDD. Alcohol-use disorders and SZ showed only small discrepancies related to the method of diagnoses. Chantarujikapong et al. (1997) compared the diagnostic utility of the DIS III and the Alcohol Dependence Questionnaire (ADS) in a sample of 143 homeless substance-abusing women, most of whom were AAs. Overall, the ADS showed acceptable agreement with the DIS in this study across substance use, personality disorder, and PTSD diagnoses. More recently, Cook et al. (2010) estimated the 12-month prevalence of psychiatric disorders and frequency of treatment seeking in 744 low-income pregnant women. 57.5 % of the sample was AA and the DIS-IV was used to establish diagnoses. AA women were less likely than Caucasians to have at least one mental disorder, and Caucasians also had a higher prevalence of lifetime psychiatric disorders, particularly affective disorders, substance-abuse disorders, and attention deficit hyperactivity disorder (ADHD).

Schedule of Affective Disorders and Schizophrenia (SADS)

The Schedule for Affective Disorders and Schizophrenia (SADS; Endicott and Spitzer 1978) is a long-standing semistructured interview with a strong presence in the literature. The SADS is established as a gold standard for the assessment of its eponymous constructs. It contains modules for current and lifetime diagnoses. The SADS should be used by an experienced clinician with a strong background in diagnostic assessment, and may take from 45 to 75 min to assess for current symptomatology, and an additional 15 min to 1 h for additional assessment of lifetime mood and psychotic episodes (Rogers 2001). Note that these times reflect administration of the modules in the order discussed.

Applications with AAs

Few of the many studies using the SADS have focused on AA clients. Some have demonstrated striking differences in diagnostic prevalence. In one such early study, Vernon et al. (1982) found that AAs and Mexican Americans had a rate of bipolar II disorder twice that of Caucasians. Other research using this measure has found that for bipolar I disorder, no such differences exist between Caucasian and AA clients, and that suicidality is rated as being far lower by the SADS for AAs than for Caucasians (Dilsaver et al. 1994). Other work has pointed out that being AA is predictive of a diagnosis of SZ using the SADS (Pavkov et al. 1989). There is a stark absence of recent comparative diagnostic research using this measure in the assessment of AA clients.

Though this measure has demonstrated great utility in the diagnosis of affective disorders and SZ spectrum disorders, it is somewhat limited in scope. Further, given the time investment necessary, this measure may be best used when a thorough characterization of a client's affective or psychotic symptoms is warranted.

Present State Examination (PSE)

The Present State Examination (PSE; Wing et al. 1967) is one of the most commonly used structured interviews within the International Classification of Diseases (ICD) framework; therefore it is more popular outside of the USA. The PSE focuses on the presence and description of any major disorder or symptoms and not on the association of specific symptoms with clinical diagnoses (Wing et al. 1967; Rogers 2001). For example, symptoms of depression may be subsumed in several diverse syndromes. The PSE is currently in its tenth edition and previous editions have varied significantly. This should be taken into account when reviewing its validity and reliability studies. The PSE 10th ed. (PSE-10; World Health Organization 1994) has total of 1224 items that evaluate from a descriptive perspective a wide variety of symptoms. Generally, the PSE has been shown to have moderately to high validity and reliability (e.g., Huxley et al. 1987; Mignolli et al. 1988; Lesage et al. 1991; Wilmink and Snijders 1989; Peveler and Faiburn 1990).

Applications with AAs

To our knowledge there are not any studies in the literature that investigate the validity of the PSE with the AA population or focuses on cultural aspects regarding AA; however the PSE has been widely used in cross-cultural research outside the USA (e.g., Hodiamont et al. 1987; Vasquez-Barquero et al. 1987). It has been translated into 40 different languages and it has been shown to be clinically useful in European countries (e.g., Garyfallos et al. 1991), English-speaking countries (e.g., Romans-Clarkson et al. 1990), and African countries (e.g., Katz et al. 1988). In a cross-national study of ten countries, results indicated that developing countries (e.g., India, Nigeria) had a higher incidence of brief psychoses than industrialized countries (e.g., Ireland, Japan, USA; Susser and Wanderling 1994). In Great Britain, people of African descent had a higher probability of being hospitalized or detained (Goater et al. 1999). In another study, Katz (1988) investigated cultural specific dimensions of psychotic disorders, in which Indian and Nigerian patients with SZ were assessed using the PSE. Indian patients showed more systematized delusions and olfactory hallucinations, whereas Nigerian patients showed more delusions of control, thought insertion, and visual hallucinations. Swartz et al. (1985) suggested that some symptoms were culturally bound and therefore difficult to interpret meaningfully. The application and validation of the PSE with diverse cultures is one of its major strengths. Nevertheless, the fact that it is based on the ICD and the lack of studies specifically focused on AAs limits the use of the PSE for US clinicians.

Schedules for Clinical Assessment of Neuropsychiatry (SCAN)

The Schedules for Clinical Assessment of Neuropsychiatry (SCAN; WHO 1994) was developed under the sponsorship of the World Health Organization (WHO) and its main component is the PSE 10th ed. The three main aims of the SCAN are rigorous clinical observation, common clinical language standardized across different diagnostic systems, and accumulation of clinical knowledge (Wing 1996). The SCAN is comprised of 27 sections. The PSE-10 (described above) is the first 25 sections. The other two sections are the Item Group Checklist (IGC) and the Clinical History Schedule (CHS). The IGC is composed of 59 ratings defined in PSE-10 terms and based on secondary sources such as records or informant. The CHS is an optional section of 88 items for the recording of childhood data, intellectual functioning, social relationships, adult personality, clinical diagnoses, and physical illness. The SCAN is a very detailed structured interview that contains standardized questions and optional probes.

Administration time is approximately 60–90 min. SCAN interviewers are required to undergo intensive training prior to administration. The SCAN was intended to be used by mental health professionals; however, it can also be administered by nonprofessionals that undergo extensive training and direct supervision (Rogers 2001). Considering that the PSE is the major component of the SCAN, their purposes are similar in that the focus is on symptomatology and the description of clinical phenomena more than linking symptoms to a diagnoses. The validity and reliability of the SCAN have been shown to be moderate to high (e.g., Easton et al. 1997; Wing et al. 1998; Brugha et al. 1999; Farmer et al. 1993, 1996; Hapke et al. 1998).

Applications with AAs

To our knowledge there are no studies in the research literature that focus on the effectiveness of the SCAN to evaluate AAs. Many of the reliability and validity studies included black people, but ethnic differences were not examined or not reported. More studies are needed to address the utility of the SCAN with AA.

Composite International Diagnostic Interview (CIDI)

The CIDI was developed under the sponsorship of the WHO. It was based on the DIS and expanded with questions from the PSE, but several items were modified to be more useful internationally. The main purpose of the CIDI was to facilitate cross-cultural epidemiologic and comparative studies. Further, it is a highly structured interview. Nonprofessionals can administer it after extensive training, and it is easily translated to different languages (Robins et al. 1988). In addition, the CIDI provides both DSM and ICD diagnoses. Administration time varies from 75 to 105 min depending on the experience of the interviewer. The current version of the CIDI (3.0) is composed of 42 sections that assess a wide variety of disorders included in the DSM-IV and ICD-10. There is a computer version available, the CIDI 3.0 Computer Assisted Personal Interview (CAPI V21.1.3). The CIDI has been translated to several languages and training materials are available in Arabic, Dutch, English, and Spanish (WHO, 2004). A CIDI Primary Health Care Version developed to address psychological problems frequently seen in medical settings (Janca et al. 1994) is also available. The reliability and validity of the CIDI has been shown to be moderate to high (Witchen et al. 1994; Peters et al. 1998; Cottler et al. 1997; Janca et al. 1992). However, these studies frequently report on data across language versions and diagnostic frameworks, so they should be interpreted with caution. Nevertheless, this is expected considering the cross-cultural focus of the CIDI and the multiple versions that are available.

Applications with AAs

Hickman III et al. (2010) conducted a study investigating past-year mental illness among 3411 adults identifying as Black Americans, with family having lived in the USA for three generations. A modified version of the computer version of the CIDI found lifetime prevalence of any mental disorder was 36.9 %, the prevalence for past-year mental disorders was 18.1 %, and the prevalence for past-month mental disorders was 4.7 %. low income and education were associated with high prevalence. Hickman III et al. also noted that divorced/separated or widowed status, and residence in the Northeast or Midwest predicted MDD for AA clients. Further studies on various samples of this database (Aranda et al. 2012; Woodward et al. 2013) have examined the prevalence of MDD in those 50–55 and older, and observed that older AAs had lower rates of MDD than whites or Caribbean black participants, that higher age and lower disability for all groups was associated with lower MDD than younger age and more disability, and that non-Hispanic whites and women had the highest prevalence of MDD across groups. In another study using the same database along with samples from two other large databases, Chou et al. (2012) found that perceived racial discrimination was associated with higher endorsement of psychiatric symptoms across racial minority groups. Notably, for AAs in this sample, higher rates of perceived racial discrimination were associated with elevated PTSD relative to Asian Americans, and had significantly higher rates of discrimination than other groups.

These studies suggest that the CIDI is a useful tool to assess AAs. However, there are very few studies that investigate the effectiveness of the CIDI compared to other interviews when assessing AAs. Green et al. (2012) conducted a study comparing DSM-IV diagnoses assessed by the CIDI and the Schedule for Affective Disorders and Schizophrenia for School Age Children (KSADS). Findings indicated that the sensitivity of the CIDI varied by race/ethnicity for ADHD, agoraphobia, panic disorder, and PTSD. Further, the specificity of the CIDI varied by race/ethnicity for agoraphobia. In line with these results Alegria et al. (2009) reported that the CIDI was problematic for accurate PTSD diagnosis in racial/ethnic minority groups. More studies like these are needed to further investigate measurement validity according to race and ethnicity.

Mini International Neuropsychiatric Interview (MINI)

As a diagnostic screening tool, the clinician-rated MINI (Sheehan et al. 1998) offers great utility and is indeed among the most widely used psychiatric structured diagnostic interviews. It is a quick, robust tool for the assessment of current, past, and lifetime diagnoses of 17 common ICD-10 and DSM disorders. Further, given its alignment with ICD-10 diagnostic categories, less may need to change for the MINI as the transition to the DSM-5 is made. In a series of validation studies (Sheehan et al. 1998), the MINI showed high negative predictive value (>.92) for all assessed diagnostic categories when compared with the SCID, and similarly (>.88) for the CIDI. This instrument also requires substantially less time to administer than either, has good-to-excellent kappa for the great majority of diagnostic criteria, and requires less training to properly administer than do many structured and semistructured interviews, such as the SCID. Two 120-min training sessions have been shown to be sufficient for raters with an undergraduate level education (Black et al. 2004; Gunter et al. 2008). It is a well-validated measure translated into more than 30 languages and is thus a good tool for the assessment of English-as-a-second-language clients.

Sheehan et al. (1997) assessed the validity of the MINI in relation to other gold standard structured diagnostic systems (SCID, CIDI), and for the majority of diagnostic categories found high agreement. This study, however, used an exceptionally homogeneous sample (96 % white), leaving much to the imagination for its validity within an AA sample.

Upon validation, the MINI demonstrated good interrater and test–retest reliability for the original sample of participants (Sheehan et al. 1998, 1997). Furthermore, good-to-great reliability coefficients were obtained in all studies noted above. Since robust demographic data were not presented, again, the reliability of this measure for the assessment of AA clients cannot be assumed.

Applications with AAs

The MINI has demonstrated success in the classification of depression and rule-out of severe psychiatric pathology in a small sample of low-income HIV-positive AA clients (Himelhoch et al. 2011). The MINI Kid has also been proven an effective screener for African child refugees in Austria (Huemer et al. 2011). There have been several studies demonstrating its effective use with Kurdish language speakers in Iraqi Kurdistan (Mitchell et al. 2011); Japanese language speakers in Japan (Otsubo et al. 2005); as a French and English socioculturally adapted version for asylum seekers in Geneva (Durieux-Paillard et al. 2006); to assess depression and depressive symptoms in a group of breast cancer patients in Lagos, Nigeria (Popoola and Adewuya 2012); and for the assessment of ADHD in an Iowa state prison sample (Westmoreland et al. 2010). The MINI is also used widely in clinical psychological and psychiatric practice around the world. Clearly then, this measure has demonstrated great generalizability across translations and cultures with at most minimal alterations (see Durieux-Paillard et al. 2006).

Though there is a good deal (comparatively) of research extolling the use of the MINI with a broad range of culturally diverse clients, there is far less information specifically validating its use with AAs. Even studies such as those addressing African refugees in Austria or Africans in Lagos are only tangentially related to such persons in the USA, and leave room for acculturative differences in the manifestation and interpretation of psychopathology.

Scale for the Assessment of Positive Symptoms (SAPS)

The Scale for the Assessment of Positive Symptoms (SAPS; Andreasen 1984) is considered a gold standard rubric for the assessment of positive symptoms in psychopathology, notably in persons with SZ. It is also used in other forms of mental illness for which positive symptoms are a prominent aspect, such as bipolar disorder. It provides four separate subscales addressing domains of positive symptomatology commonly found in persons with psychotic disorders. Given its longstanding use as a research and clinical tool, it is not hard to understand why this tool is considered by many to be a "gold standard."

An excellent treatment of the reliability, validity, internal consistency, and score differentials as gained from well-trained administrators within a focused SZ setting is given in Chap. 16 of this volume, "Assessing Psychosis in African American Clients."

Clinician Administered PTSD Scale for DSM-IV (CAPS)

The CAPS (Blake 1998) is a structured interview that aims to assess DMV-IV PTSD criteria in detail. Each inclusion criteria is evaluated according to frequency and intensity. Descriptive information is also inquired for relevant examples and to address inherent inaccuracy in self-reporting. Further, current and lifetime episodes are evaluated. Administration time is ranges from approximately 45 min to 1 h. The CAPS can be administered by nonprofessionals with a moderate level of training and extensive interviewing experience. The CAPS has been shown to have high reliability (Weathers et al. 1999; Fleming and Difede 1999) and validity (Davis et al. 2000; Weathers et al. 1999).

Applications with AAs

In two separate studies of AA and white service members (Frueh et al. 2004; Monnier et al. 2002), the CAPS did not demonstrate significant differences in anxiety, paranoia, dissociation, SZ, depression, and PTSD symptomatology between racial groups. Further, studies have assessed AAs with and without trauma exposure with the CAPS (Mellman 2009), have used the CAPS with the SCID to screen AA adults in primary care for PTSD (Graves et al. 2011), and have used the CAPS to diagnose participants with PTSD in a study examining the influence of race on CBT for PTSD (Lester 2010). The aim of the above entwined studies was not to evaluate the sensitivity of the CAPS to cultural factors or its validity in diagnosing AAs. However, compared to other interviews the CAPS has been used more often with AAs and seems to be a effective diagnostic tool for this population.

Other structured and semistructured interviews that are available for clinicians to use include the Royal Park Multidiagnostic Instrument for Psychosis (RPMIP; McGorry et al. 1988), the Diagnostic Interview for Genetic Studies (DIGS; Nurnberger et al. 1994), the Clinical Interview Schedule—Revised (CIS-R; Lewis et al. 1992), the Psychiatric Diagnostic Interview (PDI; Othmer et al. 1981), and the Polydiagnostic Interview (PODI; Phillp and Maier 1986). These clinical interviews are not as widely used as those described above and there is limited empirical information, especially concerning their use with AAs or in cross-cultural application. A review of these measures is beyond the scope of the current chapter. For a brief review of these tools and for a more extensive review of the most popular tools see Rogers (2001).

Conclusions and Recommendations

Currently there is a paucity of research regarding the validity and accuracy of the aforementioned structured and semistructured interviews when used with minority populations, in particular AAs. Some of the instruments have been used with AA populations; however, there are very few studies that focus specifically on the cross-cultural validity and reliability of the instrument with AAs. Nevertheless, the structured and semistructured interviews reviewed in this chapter could potentially be helpful in reducing diagnostic bias due to cultural factors (criterion and information variance). For example, structured and semistructured interviews ensure that a wide variety of symptoms are covered, standardize language, reduce the possibility of clinician stereotypes interfering because they restrict the range of questions, and also reduce the possibility of misinterpreting answers because many of the answers tend to be dichotomized. Several of the interviews reviewed seem to be very promising instruments that could be applied cross-culturally. Instruments such as the CIDI and the MINI were developed for cross-cultural applications. With the amount of research available it is not possible to conclude which interview would be the best to use with AAs. However, most of the interviews presented have been used in studies that included AAs. It is difficult to draw clear conclusions from these studies due to the substantial differences in the samples used, instruments used, and also the version of the instrument used. Nevertheless, some of the structured and semistructured interviews that have been used more often with AAs are the SCID, the DIS, the CAPS, the CIDI, and the SAPS. On the other hand, because some of these interviews are so standardized (fully structured), they cannot be tailored to the particular cultural needs of a population and questions cannot be modified accordingly. In the case of PTSD, it has been suggested that fully structured interviews (e.g., CIDI) are differentially biased towards minorities because they are less able than semistructured interviews to take the cultural context of trauma and trauma-related symptoms into account (Alarcon 1995). However, when using a semistructured instrument the clinician has to be culturally component enough to be able to interpret the cultural context correctly. Considering the lack of research in this area, clinicians' awareness of possible diagnostic bias and cultural competence seem to be essential to avoid diagnostic errors. If a clinician is culturally incompetent a fully structured interview, such as the CIDI, would be more appropriate to use in order to reduce as much variance as possible. Nevertheless, a culturally incompetent clinician should attempt to refer a minority client to somebody that is an expert in that client's culture. If a clinician is culturally competent a semistructured interview, such as the SCID, that allows for the incorporation of questions about cultural and ethnic values and factors would be recommended. For example, these questions could be added during the SCID overview. Another possibility is using both a structured interview and an interview that addresses cultural and ethnic factors. Grieger and Ponterotto (1995) developed a useful conceptual framework that can be integrated to the interview process or used as a complimentary assessment with minority group members, including AAs. Based on the constructs of worldview and acculturation they developed the following components as being fundamental in culturally situating the client and conceptualizing the client's presenting problems within a relevant cultural context: the client's level of psychological mindedness; the family's level of psychological mindedness; the client's and family's attitudes towards helping and counseling; the client's level of acculturation; the family's level of acculturation; and the family's attitudes toward acculturation. Grieger (2008) expanded this conceptual model by developing 11 categories: (1) Problem Conceptualization and Attitudes Towards Helping, (2) Cultural Identity, (3) Level of Acculturation, (4) Family Structure and Expectations, (5) Level of Racial/Cultural Identity Development, (6) Experiences with Bias, (7) Immigration Issues, (8) Existential/Spiritual Issues, (9) Counselor Characteristics and Behaviors, (10) Implications of Cultural Factors Between the Counselor and the Client, (11) Summary of Cultural Factors and Implications for Diagnosis, Case Conceptualization, and Treatment. Using a conceptual framework such as this can be helpful in acknowledging and identifying cultural and ethnic factors that can be influencing the diagnostic interview process with an AA client. Grieger (2008) presented a Cultural Assessment Interview Protocol, which contains a set of questions based on her conceptual framework. Questions such as these could potentially be integrated in a semistructured interview or asked separately (See Grieger 2008 for details) The DSM-5 also includes an interview to aid with the cultural formulation that contains several useful questions.

Though semistructured interviews do indeed provide a good scaffold to support empirically based diagnostic decision making, they are just that: a scaffold. They do not remove the burden of well-honed objective clinical skill from the equation. It is imperative that a clinician understands his or her biases, the background of a patient, and does not rely on the instrument to provide context for a unique person. Further, it is imperative that a clinician realize the extent to which many diagnostic considerations are indeed simply a measure of degree rather than always objective and self-evident. Considering the research reviewed in this chapter, when interviewing AAs cultural competence is essential. For AAs, the cultural sensitivity of the mental health provider is one of the most important characteristics. Culturally sensitive counselors recognize that race or culture might play role in the patient's problem. Culturally blind counselors tend to focus on aspects other than race when dealing with the presenting problem. Counselors that are culturally sensitive are seen as more competent by AAs than those who are culture blind (Pomales et al. 1986; Want et al. 2004). Cultural competence involves more than just awareness and practice of the considerations reviewed in this chapter. It is an ongoing process that includes formal training, self-education, consultations with colleagues that are more knowledgeable, use of collateral information when conceptualizing cases, and attending workshops and continuing education programs. U.S. Department of Health and Human Services. (2001) In addition, experience working with AA clients and immersion to experience the AA culture firsthand are fundamental.

References

Adebimpe, V. R. (1981). Overview: White norms and psychiatric diagnosis of Black patients. *American Journal of Psychiatry,138*(3), 279–285.

Akan, G. E., & Grilo, C. M. (1995). Sociocultural influences on eating attitudes and behaviors, body image, and psychological functioning: A comparison of African-American, Asian-American, and Caucasian college women. *International Journal of Eating Disorders,18*(2), 181–187.

Aklin, W. M. & Turner, S. M. (2006). Toward understanding ethnic and cultural factors in the interviewing process. *Psychotherapy: Theory, Research, Practice, Training,43*(1), 50–64.

Al-issa, I. (1995). The illusion of reality or the reality of illusion: Hallucinations and culture. *British Journal of Psychiatry,166,* 368–373.

Alarcon R.D. (1995) Culture and psychiatric diagnosis. Impact on DSM-IV and ICD-10. *Psychiatric Clinics of North America, 18*(3), 449–465.

Alterman, H., Snider, E. C., Cacciola, J S., Brown L. S., Jr., Zaballero, A. & Sidiqui, N. (1996). Evidence of response set effects in structured research interviews. *Journal of Nervous and Mental Disease,184,* 403–410.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders fourth edition, text revision*. Washington, D.C.: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders fifth edition*. Arlington: American Psychiatric Association.

Andreasen, N. C. (1984). *Scale for the assessment of positive symptoms*. Iowa City: University of Iowa Press.

Aranda, M. P., David H. Chae, Karen D. Lincoln, Robert Joseph Taylor, Amanda Toler Woodward4 and Linda M. Chatters. (2012). Demographic correlates of DSM-IV major depressive disorder among older African Americans, Black Caribbeans, and non-Hispanic Whites: results from the National Survey of American Life. *IntERNATIONAL Journal of Geriatric Psychiatry; 27*: 940–947.

Alegria M., Shrout P.E., TorresM., Lewis-Fernandez R., Abelson J.M., Powell M., Interian A., Lin J., Laderman M., Canino G. (2009) Lessons learned from the clinical reappraisal study of the Composite International Diagnostic Interview with Latinos. *International Journal of Methods in Psychiatric Research, 18*(2), 84–95.

Bae, S. W., & Brekke, J.S. (2002). Characteristics of Korean-Americans with schizophrenia: A cross-ethnic comparison with African-Americans, Latinos, and Euro-Americans. *Schizophrenia Bulletin,28*(4), 703–717.

Barnes, A. (2008). Race and hospital diagnoses of schizophrenia and mood disorders. *Social Work,53,* 77–83.

Bauer, S. M., Schanda, H., Karakula, H., Olajossy-Hilkesberger, L., Rudaleviciene, P., Okribelashvili, N., Chaudhry, H. R., Idemudia, S. E., Gscheider, S., Ritter, K., & Stompe, T. (2011). Culture and the prevalence of hallucinations in schizophrenia. *Comprehensive Psychiatry,52*(3), 319–325.

Bell, C., & Mehta, H. (1980). The misdiagnosis of black patients with manic depressive illness. *Journal of the National Medical Association,72,* 141–145.

Bentacourt, H., & Lopez, S. R. (1993). The study of culture, ethnicity, and race in American psychology. *American Psychologist,48,* 629–637.

Blake, D. D., Weathers, F., Nagy, L. M., Kaloupek, D. G., Charney, D. S., & Keane, T. M. (1998). Clinician Administered PTSD Scale for DSM-IV. Boston: National Center for Posttraumatic Stress Disorder.

Black, D. W., Arndt, S., Hale, N., & Rogerson, R. (2004). Use of the Mini International Neuropsychiatric Interview (MINI) as a screening tool in prisons: Results of a preliminary study. *Journal of the American Academy of Psychiatry and the Law Online,32*(2), 158–162.

Bresnahan, M., Begg, M. D., Brown, A., Schaefer, C., Sohler, N., Insel, B., Vella, L., Susser, E. (2007). Race and risk of schizophrenia in a US birth cohort: Another example of health disparity? *International Journal of Epidemiology, 36,* 751–758.

Breslau, J., Aguilar-Gaxiola, S., Kendler, K. S., Su, M., Williams, D., & Kessler, R. C. (2006). Specifying race-ethnic differences in risk for psychiatric disorder in a USA national sample. *Psychological Medicine, 36*(1), 57–68.

Brown, C., Schulberg, H. C., & Madonia, M. J. (1996). Clinical presentations of major depression by African Americans and whites in primary medical care practice. *Journal of Affective Disorders, 41,* 181–191.

Brown, C., Schulberg, H. C., Sacco, D., Perel, J. M., & Houck, P. R. (1999). Effectiveness of treatments for major depression in primary medical care practice: A post hoc analysis of outcomes for African American and White patients. *Journal of Affective Disorders, 53*, 185–192.

Brugha, T. S.; Bebbington, P. E.; Jenkins, R.; Meltzer, H.; Taub, N. A.; Janas, M.; Vernon, J. (1999). Cross validation of a general population survey diagnostic interview: A comparison of CIS-R with SCAN ICD-10 diagnostic categories. *Psychological Medicine, 29*(5), 1029–1042.

Carlson, D. K. (2004). Racial profiling seen as pervasive, unjust. Retrieved on January 3rd 2014 from: www.gallup.com/poll/12406/racial-profiling-seen-pervasive-unjust.aspx.

Chantarujikapong, S. I., Smith, E. M., & Fox, L. W. (1997). Comparison of the Alcohol Dependence Scale and Diagnostic Interview Schedule in homeless women. *Alcoholism: Clinical and Experimental Research,21*(4), 586–595.

Charters, L. M., Taylor, R. J., Jackson, J. S., & Lincoln, K. D. (2008). Religious coping among African Americans, Caribbean Blacks and non-Hispanic Whites. *Journal of Community Psychology, 36*(3), 371–386.

Chou, T., Anu Asnaani, & Stefan G. Hofmann. (2012). Perception of Racial Discrimination and Psychopathology Across Three U.S. Ethnic Minority Groups. *Cultural Diversity and Ethnic Minority Psychology, 18*(1), 74–81.

Columbia DISC Development Group. (1999). National Institute of Mental Health Diagnostic Interview Schedule for Children (NIMH-DISC). Unpublished report, Columbia University/New York State Psychiatric Institute.

Constantine, M. G. (2007). Racial microaggressions against African American clients in cross-racial counseling relationships. *Journal of Counseling Psychology,54*(1), 1–16.

Cook, C. A. L., Flick, L. H., Homan, S. M., Campbell, C., McSweeney, M., & Gallagher, M. E. (2010). Psychiatric disorders and treatment in low-income pregnant women. *Journal of Women's Health,19*(7), 1251–1262.

Cottler, Linda B.; Grant, Bridget F.; Blaine, Jack; Mavreas, Vanetsanos; Pull, Charles; Hasin, Deborah; Compton, Wilson M.; Rubio-Stipec, Maritza; Mager, Doug. (1997). Concordance of the DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. *Drug and Alcohol Dependence, 47*(3), 195–205.

Cottler, L. B., Compton, W. M., Ridenour, A., Abdahlla, A. B., & Gallagher, T. (1998). Reliability of self-reported antisocial personality disorder symptoms among substance abusers. *Drug and Alcohol Dependence, 49,* 189–199.

Cross, W. E. (1995). The psychology of Nigrescence: Revising the Cross model. In J. G. Ponterotto, J. M. Casas, L. M. Suzuki, & C. M. Alexander (Eds.). Handbook of multicultural counseling (pp. 93–122). Thousand Oaks, CA: Sage.

Dana, R. H. (2008). Clinical diagnosis in multicultural populations. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (pp. 107–131). San Francisco: Wiley.

Darley, J. & Fazio, R. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35*, 867–881.

Devine, P. G., & Baker, D. M. (1991). Measurement of racial stereotype subtyping. *Personality and Social Psychology Bulletin, 17*, 44–50.

Davis, L.L., Nugent, Alexandria L.; Murray, Jenny; Kramer, Gerald L.; Petty, Frederick. (2000). Nefazodone treatment for chronic posttraumatic stress disorder: An open trial. *Journal of Clinical Psychopharmacology, 20* (2), 159–164.

DeCoux Hampton, M. (2007). The role of treatment setting and high acuity in the overdiagnosis of schizophrenia in African Americans. *Archives of Psychiatric Nursing,21*(6), 327–335.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18.

Dijksterhuis, A. & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology, 33*, 1–40.

Dilsaver, S. C., Chen, Y. W., Swann, A. C., Shoaib, A. M., & Krajewski, K. J. (1994). Suicidality in patients with pure and depressive mania. *American Journal of Psychiatry,151*(9), 1312–1315.

Diala, C. C., Muntaner, C., Walrath, C., Nickerson, K., LaVeist, T.,&Leaf, P. (2001). Racial/ethnic differences in attitudes toward seeking professional mental health services. *American Journal of Public Health, 91*(5), 805–807.

Durieux-Paillard, S., Whitaker-Clinc, B., Bovier, P. A., Eytan, A. (2006). Screening for major depression and posttraumatic stress disorder among asylum seekers: Adapting a standardized instrument to the social and cultural context. *Canadian Journal of Psychiatry, 51,* 587–597.

Eack, S. M., Bahorik, A. L., Newhill, C. E., Neighbors, H. W., & Davis, L. E. (2012). Interviewer-perceived honesty as a mediator of racial disparities in the diagnosis of schizophrenia. *Psychiatric Services,63*(9), 875–880.

Easton, C., Meza, E., Mager, D., Ulüg, B., Kilic, C., Gögüs, A., & Babor, T. F. (1997). Test-retest reliability of the alcohol and drug use disorder sections of the schedules for clinical assessment in neuropsychiatry (SCAN). *Drug and Alcohol Dependence, 47*(3), 187–194.

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry,35*(7), 837.

Escobar, J. I. (2012). Diagnostic Bias: Racial and Cultural Issues. *Psychiatric Services, 63*(9), 847.

Farmer, Anne; Cosyns, Paul; Leboyer, Marion; Maier, Wolfgang; Mors, Ole; Sargeant, Matthew; Bebbington, Paul; McGuffin, Peter. (1993). A SCAN-SADS comparison study of psychotic subjects and their first-degree relatives. *European Archives of Psychiatry and Clinical Neuroscience, 242*(6), 352–356.

Farmer, Anne; Chubb, Helene; Jones, Irene; Hillier, Janis; Smith, A.; Borysiewicz, L. (1996) Screening for psychiatric morbidity in subjects presenting with chronic fatigue syndrome. *The British Journal of Psychiatry, 168*(3), 354–358.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W. (Nov 2002). Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute.

Fleming, M. P., & Difede, J. (1999). Effects of varying scoring rules of the Clinician Administered PTSD Scale (CAPS) for the diagnosis of PTSD after acute burn injury. *Journal of Traumatic Stress, 12*(3), 535–542.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method of grading cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Frueh, B. C., Brady, K. L., & de Arellano, M. A. (1998). Racial differences in combat-related PTSD: Empirical findings and conceptual issues. Clinical Psychology Review, 18, 287–305.

Frueh, B. C., Elhai, J. D., Monnier, J., Hamner, M. B., Knapp, R. G. (2004). Symptom patterns and service use among African American and Caucasian veterans with combat-related PTSD. *Psychological Services,1(1)*, 22–30.

Fry, R. (2010). Hispanics high school dropouts and the GED. Retrieved on January 3rd, 2014 from: http://www.pewhispanic.org/2010/05/13/ii-high-school-dropouts-and-geds/.

Garb, H. N. (1997). Race bias, social class bias, and gender bias in clinical judgment. *Clinical Psychology: Science and Practice, 4,* 99–120.

Garyfallos, G., Karastergiou, A., Adamapoulou, A., Moutzoukis, C., Alagiozidou, E., Mala, D., & Garyfallos, A. (1991). Greek version of the General Health Questionnaire: Accuracy of translation and validity. *Acta Psychiatrica Scandinavica, 84,* 371–378.

Goater, N., King, M., Cole, E., Leavey, G., Johnson-Sabine, E., Blizard, R., & Hoar, A. (1999). Ethnicity and outcome of psychosis. *British Journal of Psychiatry, 175,* 34–42.

Graves, R. E., Freedy, J. R., Aigbogun, N. U., Lawson, W. B., Mellman, T. A., Alim, T. N. (2011). PTSD treatment of African American adults in primary care: The gap between current practice and evidence-based treatment guidelines. *Journal of the National Medical Association,103*(7), 585–593.

Green, B. L., Grace, M. C., Lindy, J. D., & Leonard, A. C. (1990). Race differences in response to combat stress. *Journal of Traumatic Stress, 3*(3), 379–393.

Green, J. G., MICHAEL J. GRUBER, RONALD C. KESSLER, JULIA Y. LIN, KATIE A. MCLAUGHLIN, NANCY A. SAMPSON, ALAN M. ZASLAVSKY & MARGARITA ALEGRIA. (2012). Diagnostic validity across racial and ethnic groups in the assessment of adolescent DSM-IV disorders. *International Journal of Methods in Psychiatric Research, 21*(4), 311–320.

Grieger, I. & Ponterotto, J. G. (1995). A framework for assessment in multicultural counseling. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki & C. M. Alexander (Eds.). *Handbook of multicultural counseling (pp. 357–374)*. Thousand Oaks, CA: Sage.

Grieger, I. (2008). A cultural assessment framework and interview protocol. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (pp. 132–161). San Francisco: Wiley.

Gunter, T. D., Arndt, S., Wenman, G., Allen, J., Loveless, P., Sieleni, B., & Black, D. W. (2008). Frequency of mental and addictive disorders among 320 men and women entering the Iowa prison system: use of the MINI-Plus. *Journal of the American Academy of Psychiatry and the Law Online,36*(1), 27–34.

Gushue, G. V. (2004). Race, color-blind racial attitudes, and judgments about mental health: A shifting standards perspective. *Journal of Counseling Psychology,51,* 398–407.

Hapke, Ulfert; Rumpf, H.-J.; John, U. (1998). Differences between hospital patients with alcohol problems referred for counselling by physicians' routine clinical practice versus screening questionnaires. *Addiction, 93*(12), 1777–1785.

Helzer, J.E., Robins, L.N., McEnvoy, L. T., Spitznagel, L. M., Stoltzman, R. K., Farmer, A., & Brockington, I. F. (1985). A comparison of clinical and diagnostic interview schedule diagnoses. *Archives of General Psychiatry, 42,* 657–666.

Hesselbrock, V., Stabenau, J., Hesselbrock, M., Mirkin, P., & Meyer, R. (1982). A comparison of two interview schedules: The Survey for Affective Disorders and Schizophrenia-Lifetime and National Institute of Mental Health Diagnostic Interview Schedule. *Archives of General Psychiatry, 39,* 674–677.

Hickman III, N. V., Kevin L. Delucchi, & Judith J. Prochaska (2010) . A Population-Based Examination of Cigarette Smoking and Mental Illness in Black Americans. *Nicotine & Tobacco Research, 12*(11), 1125–1132.

Himelhoch, S., Mohr, D., Maxfield, J., Clayton, S., Weber, E., Medoff, D., & Dixon, L. (2011). Feasibility of telephone-based cognitive behavioral therapy targeting major depression among urban dwelling African-American people with co-occurring HIV. *Psychology, Health & Medicine,16*(2), 156–165.

Hodiamont, P., Peer, N., & Sybern, N. (1987). Epidemiological aspects of psychiatric disorders in Dutch health area. *Psychological Medicine, 17,* 495–505.

Holmes, S. A., & Morin, R. (2006). Black men torn between promise and doubt. Retrieved on January 1st from: http://www.africanamerica.org/topic/black-men-torn-between-promise-and-doubt.

Huemer, J., Niranjan Karnik . Sabine Voelkl-Kernstock, Elisabeth Granditsch, Belinda Plattner, Max Friedrich, Hans Steiner. (2011). Psychopathology in African Unaccompanied Refugee Minors in Austria. *Child Psychiatry and Human Development, 42*, 307–319.

Huxley, P., Korer, J., & Tolley, S. (1987). The psychiatric "caseness" of clients referred to an urban social services department. *British Journal of Sociology, 17,* 507–520.

Hyers, L. L. (2001). A secondary survey analysis study of African American ethnic identity orientation in two national samples. *Journal of Black Psychology, 27*, 139–171.

Janca A, Helzer J. (1990). DSM-III-R Criteria Checklist. *Diagnostic Interview Schedule Newsletter*, *7*, u17.

Janca, A.; Robins, L. N.; Bucholz, K. K.; Early, T. S.; Shayka, J. J.(1992). Comparison of Composite International Diagnostic Interview and clinical DSM-III—R criteria checklist diagnoses. *Acta Psychiatrica Scandinavica, 85*(6), 440–443.

Janca, A.; Üstün, T. B.; (1994). Sartorius, N. New versions of World Health Organization instruments for the assessment of mental disorders. *Acta Psychiatrica Scandinavica, 90*(2), 73–83.

Jeste, D. V., Lindamer, L. A., Evans, J., & Lacro, J. P. (1996). Relationship of ethnicity and gender to schizophrenia andpharmacology of neuroleptics. *Psychopharmacology Bulletin,32*(2), 243–251.

Jimenez, D. E., Alegría, M., Chen, C., Chan, D., & Laderman, M. (2010). Prevalence of psychiatric illnesses in older ethnic minority adults. Journal of the American Geriatrics Society, 58(2), 256–264.

Jonas, B. S., Brody, D., Roper, M., & Narrow, W. E. (2003). Prevalence of mood disorders in a national sample of young American adults. *Social Psychiatry and Psychiatric Epidemiology, 38,* 618–624.

Katz, M. M., Marsalla, A., Dube, K. C., Olatawura, M., Takahashi, R., Nakane, Y., Wynne, L. C., Gift, T., Brennan, J., Sartorius, N., & Jablenski, A. (1988). On the expression of psychosis in different cultures: Schizophrenia in an Indian and in a Nigerian Community. *Culture, Medicine, and Psychiatry, 12,* 331–355.

Kales, H. C. (2000). Race, psychiatric diagnosis, and mental health care utilization in older patients. *American Journal of Geriatric Psychiatry, 8,* 301–309.

Kovess, V., & Fournier, L. (1990). The DISSA: An abridged self-administered version of the DIS: Approach by episode. *Social Psychiatry and Psychiatric Epidemiology,25*(4), 179–186.

Kulka, R. A., Schlenger, W. A., Fairbanks, J. A., Hough, R. L., Jordan, B. K., Marmar, C. R., & Weiss, D. S. (1990). *Trauma and the Vietnam War generation: Report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel.

Lester, K., Resick, P. A., Young-Xu, Y., Artz, C. (2010). Impact of race on early treatment termination and outcomes in posttraumatic stress disorder treatment.

Journal of Consulting and Clinical Psychology,78(4), 480–489.

Lesage, A. D., Cyr, M., & Toupin, J. (1991). Reliable use of the Present State Examination by psychiatric nurses for clinical studies of psychotic and nonpsychotic patients. *Acta Psychiatrica Scandinavia, 83,* 121–124.

Lewis, Glyn; Pelosi, Anthony J.; Araya, Ricardo; Dunn, Graham (1992). Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine, 22*(2),465–486.

Lewis-Fernández, R., Hinton, D. E., Laria, A. J., Patterson, E. H., Hofmann, S. G., Craske, M. G., Stein, D. J., Asnaani, A., & Liao, B. (2010). Culture and the anxiety disorders: Recommendations for DSM-V. *Depression and Anxiety,27*(2), 212–229.

Lindsey, K. P., & Paul, G. L. (1989). Involuntary commitments to public mental institutions: Issues involving the overrepresentation of blacks and the assessment of relevant functioning. *Psychology Bulletin, 106,* 171–183.

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Interrater reliability of the structured clinical interview for DSM-IV axis I disorders (SCID I) and axis II disorders (SCID II). *Clinical Psychology & Psychotherapy,18*(1), 75–79.

Lonigan, C. J., Shannon, M. P., Finch, A. J., Daugherty, T. K., & Taylor, C. M. (1991). Children's reactions to a natural disaster: Symptom severity and degree of exposure. *Advances in Behavior Research & Therapy, 13*(3), 135–154.

Mathews, C. A., Glidden, D., & Hargreaves, W. A. (2002). The effect on diagnostic rates of assigning patients to ethnically focused inpatient psychiatric units. *Psychiatric Services*, *53*(7), 823–829.

Mellman, T. A.; Brown, D. D.; Jenifer, E. S.; Hipolito, M. M. S.; & Randall, O. S. (2009). Posttraumatic stress disorder and nocturnal blood pressure dipping in young adult African Americans. *Psychosomatic Medicine, 71*(6), 627–630.

McGuire, T. G., & Miranda, J. (2008). New evidence regarding racial and ethnic disparities in mental health: Policy implications. *Health Affairs*, *27*, 393–403.

McGorry, P., Kaplan, I., Dossetor, C., Copolov, D., & Singh, B. (1988). The Royal Park Multidiagnostic Instrument for Psychosis: A comprehensive assessment for the acute psychotic episode (RPMIP). Melbourne, Australia: Department of Psychological Medicine, Monash University.

Mignolli, G., Faccinicani, C., Turit, L., Gavioli, I., & Micciolo, R. (1988). Interrater reliability of the PSE-9 (full version): An Italian study. *Social Psychiatry and Psychiatric Epidemiology*, *23*, 30–35.

Miller, S. T., Seib, H. M., & Dennie, S. P. (2001). African American perspectives on health care: The voice of the community. *Journal of Ambulatory Care Management, 24*, 37–42.

Miller, P. R., Dasher, R., Collins, R., Griffiths, P., & Brown, F. (2001). Inpatient diagnostic assessments: 1. Accuracy of structured vs. unstructured interviews. *Psychiatry Research, 105*, 255–264.

Minsky, S., Vega, W., Miskimen, T., et al. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. *Archives of General Psychiatry*, *60*, 637–644.

Mitchell, A. J., Smith, A. B., Al-Salihy, Z., Rahim, T. A., Mahmud, M. Q., & Muhyaldin, A. S. (2011). Redefining diagnostic symptoms of depression using Rasch analysis: testing an item bank suitable for DSM-V and computer adaptive testing. *Australian and New Zealand Journal of Psychiatry,45*(10), 846–852.

Monnier, J., Elhai J. D., Frueh, BC, Sauvageot J. A., Magruder, K. M. (2002). Replication and expansion of findings related to racial differences in veterans with combat-related PTSD. *Depression and Anxiety, 16*(2), 64–70.

Monteith, M. J. (1996). Contemporary forms of prejudice-related conflict: In search of a nutshell. *Personality and Social Psychology Bulletin, 22*, 461–473.

Myers, H. F., Lesser, I., Rodriguez, N., Mira, C. B., Hwang, W. C., Camp, C., et al. (2002). Ethnic differences in clinical presentation of depression in adult women. *Cultural Diversity & Ethnic Minority Psychology,8*, 138–156.

National Institute of Mental Health. (1991). NIMH Diagnostic Interview for Children, Version 2.3 Rockville: National Institute of Mental Health.

National Alliance of Mental Health (NAMI). (2004). African American Community Mental Health Fact Sheet NAMI Multicultural Action Center. Retrieved on January 1st from: http://www.nami.org/Content/ContentGroups/MIO/Fact_Sheets1/AA_MH_Disparities_04.pdf.

Neighbors, H. W., Jackson, J. S., Campbell, L., & Williams, D. (1989). The influence of racial factors on psychiatric diagnosis: A review and suggestions for research. *Community Mental Health Journal, 25,* 301–309.

Neighbors, H. W., Trierweiler, S. J., Munday, C., Thompson, E. E., Jackson, J. S., Binion, V. J., & Gomez, J. (1999). Psychiatric diagnosis of African Americans: Diagnostic divergence in clinician-structured and semi structured interviewing conditions. *Journal of the National Medical Association,91*(11), 601–612.

Neighbors, H. W., Trierweiler, S. J., Ford, B., & Muroff, J. R. (2003). Racial differences in diagnosis using a semi-structured instrument: The importance of clinical judgment. *Journal of Health and Social Behavior, 44,* 237–256.

Nickerson, K. J., Helms, J. E., & Terrell, F. (1994). Cultural mistrust, opinions about mental illness, and Black students' attitudes toward seeking psychological help from White counselors. *Journal of Counseling Psychology,41*(3), 378.

Nisbet, R. E. & Wilson, T. D. (1997). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.

North, C. S., Pollio, D. E., Thompson, S. J., Ricci, D. A., Smith, E. M., & Spitznagel, E. L. (1997). A comparison of clinical and structured interview diagnoses in a homeless mental health clinic. *Community Mental Health Journal,33*(6), 531–543.

Norris, F. H. (1992). Epidemiology of trauma: Frequency and impact of different potentially traumatic events on different demographic groups. *Journal of Consulting and Clinical Psychology, 60*(3), 409–418.

Nurnberger, John I.; Blehar, Mary C.; Kaufmann, Charles A.; York-Cooler, Carolyn; Simpson, Sylvia G.; Harkavy-Friedman, Jill; Severe, Joanne B.; Malaspina, Dolores; Reich, Theodore (1994). Diagnostic Interview for Genetic Studies: Rationale, unique features, and training. *Archives of General Psychiatry, 51*(11), 849–859.

Otsubo, T., Tanaka, K., Koda, R., Shinoda, J., Sano, N., Tanaka, S., & Kamijima, K. (2005). Reliability and validity of Japanese version of the mini-international neuropsychiatric interview. *Psychiatry and Clinical Neurosciences,59*(5), 517–526.

Othmer, E., Penick, E. C., & Powell, B. J. (1981). *The Psychiatric Diagnostic Interview*. Los Angeles: Western Psychological Services.

Oquendo, M. A., Ellis, S. P., Greenwald, S., Malone, K. M., Weissman, M. M., & Mann, J. J. (2001). Ethnic and sex differences in suicide rates relative to major depression in the United States. *American Journal of Psychiatry, 158*(10), 1652–1658.

Overall, J. E., & Gorham, D.R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports,10,* 799–812.

Paniagua, F. A. (2001). Culture-bound syndromes, cultural variations, and psychopathology. In I. Cuellar & F. A. Paniagua (Eds.), *Handbook of multicultural mental health* (pp. 139–169). San Diego: Academic Press.

Parham, T. A., & Helms, J. E. (1981). The influence of black students' racial attitudes on preferences for counselor's race. *Journal of Counseling Psychology, 28*, 250–257.

Pavkov, T. W., Lewis, D. A., & Lyons, J. S. (1989). Psychiatric diagnoses and racial bias: An empirical investigation. *Professional Psychology: Research and Practice,20*(6), 364.

Peters, Lorna; Clark, D.; Carroll, F. (1998). Are computerized interviews equivalent to human interviewers? CIDI-Auto versus CIDI in anxiety and depressive disorders. *Psychological Medicine, 28*(4), 893–901.

Peveler, R. C., & Faiburn, C. G. (1990). Measurement of neurotic symptoms by self-report questionnaire: Validity of the SCL-90R. *Psychological Medicine, 20,* 873–879.

Pew Research Center. (2010). Blacks upbeat about Black progress, prospects: A year after Obama's election. Retrieved on January 1st from: http://www.people-press.org/2010/01/12/blacks-upbeat-about-black-progress-prospects/.

Philipp, M., & Maier, W. (1986). The polydiagnostic interview: A structured interview for the polydiagnostic classification of psychiatric patients. *Psychopathology, 19*(4), 175–185.

Popoola, A. O., & Adewuya, A. O. (2012). Prevalence and correlates of depressive disorders in outpatients with breast cancer in Lagos, Nigeria. *Psycho-Oncology, 21*(6), 675–679.

Pomales, J., Claiborn, C. D., & LaFromboise, T. D. (1986). Effects of black students' racial identity on perceptions of white counselors varying in cultural sensitivity. Journal of Counseling Psychology, 34, 123–131.

Regier, D. A., Myers, J. K., Kramer, M., Robins, L. N., Blazer, D. G., Hough, R. L., Eaton, W. W., & Locke, B. Z. (1984). The NIMH Epidemiological Catchment Area Program: Historical context, major objectives, and study population characteristics. *Archives of General Psychiatry,41*(10), 934–941.

Ridley, C. A. (1984). Clinical treatment of the nondisclosing Black client: A therapeutic paradox. American Psychologist, 39, 1234–1244.

Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E., Burke, J. D., & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41*(10), 949–958.

Robins, Lee N.; Wing, John; Wittchen, Hans U.; Helzer, John E.; Babor, Thomas F.; Burke, Jay; Farmer, Anne; Jablenski, Assen; Pickens, Roy; Regier, Darrel A.; Sartorius, Norman; Towle, Leland H. (1988). The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry, 45*(12), 1069–1077.

Robins, L. N., & Regier, D. A. (Eds.). (1991). *Psychiatric disorders in America*. New York: Free Press.

Robins, L. N., Cottler, L. B., & Keating, S. (1991). *NIMH diagnostic interview schedule, version III-revised (DIS-III-R): Question by question specifications*. St. Louis: Washington School of Medicine.

Robins, L. N., Cottler, L. B., Bucholz, K., & Compton, W. (1995). *NIMH diagnostic interview schedule, version IV*. St. Louis: Washington School of Medicine.

Rogers, R. (2001). *Handbook of diagnostic and structured interviewing*. New York: Guilford Press.

Rollman,B. L.,Hanusa, B.H., Belnap, B.H.,Gardner,W., Cooper, L. A., & Schulberg, H. C. (2002). Race, quality of depression care, and recovery from major depression in a primary care setting. General Hospital Psychiatry, 24(6), 381–390.

Romans-Clarkson, S. E., Walton, V. A., Herbison, G. P., & Mullen, P. E. (1990). Psychiatric morbidity among women in urban and rural New Zealand: Psychosocial correlates. *British Journal of Psychiatry, 156,* 84–91.

Shear, M. K., Greeno, C., Kang, J., Ludewig, D., Frank, E., Swartz, H. A., & Hanekamp, M. (2000). Diagnosis of nonpsychotic patients in community clinics. *American Journal of Psychiatry,157*(4), 581–587.

Sheehan, D. V., Lecrubier, Y., Harnett Sheehan, K., Janavs, J., Weiller, E., Keskiner, A., et al. (1997). The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*, *12*(5), 232–241.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-

International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59,* 22–33.

Snowden, K. R., & Cheung, F. K. (1990). Use of inpatient mental health services by members of ethnic minority groups. *American Psychologist, 45,* 347–355.

Steiner, J. L., Tebes, J. K., Sledge, W. H., & Walker, M. L. (1995). A comparison of the structured clinical interview for DSM-III-R and clinical diagnoses. *The Journal of Nervous and Mental Disease,183*(6), 365–369.

Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797-811.

Strakowski S. M., Flaum M., Amador X., Bracha, H. S., Pandurangi, A. K., Robinson, D., & Tohen, M. (1996). Racial differences in the diagnosis of psychosis. *Schizophrenia Research,21*(2), 117–124.

Strakowski, S. M., Hawkins, J. M., Keck Jr., P. E., Mcelroy, S. L., West, S. A., Bourne, M. L., Sax, K. W., & Tugrul, K. C. (1997). The effects of race and information variance on disagreement between psychiatric emergency service and research diagnoses in first-episode psychosis. *Journal of Clinical Psychiatry*, *58*(10), 457–463.

Strakowski, S. M., Keck Jr., P. E., Arnold, L. M., Collins, J., Wilson, R. M., Fleck, D. E., et al. (2003). Ethnicity and diagnosis in patients with affective disorders. *Journal of Clinical Psychiatry,64*(7), 747–754.

Susser, E., & Wanderling, J. (1994). Epidemiology of nonaffective acute remitting psychosis vs. schizophrenia. *Archives of General Psychiatry, 51,* 294–301.

Sue, D.W., & Sue, D. (2013). *Counseling the culturally diverse: Theory and practice*. New Jersey: Wiley & Sons.

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: implications for clinical practice. American psychologist, 62(4), 271. doi: 10.1037/0003-066X.62.4.271.

Sutker, P. B., Davis, J. M., Uddo, M.., & Ditta, S. R. (1995). Assessment of psychological distress in Persian Gulf troops: Ethnicity and gender comparisons. *Journal of Personality Assessment, 64*(3), 415–427.

Swartz, L., Ben-Arie, O., & Teggin, A. F. (1985). Subcultural delusions and hallucinations: Comments on the Present State Examination in a multicultural context. *British Journal of Psychiatry, 146,* 391–394.

Taylor, R. J., Chatters, L. M., Hardison, C. B., & Riley, A. (2001). Informal social support networks and subjective well-being among African Americans. *Journal of Black Psychology, 27*(4), 439–463.

Terrell, F., & Terrell, S. (1984). Race of counselor, client sex, cultural mistrust level, and premature termination from counseling among Black clients. *Journal of Counseling Psychology, 31,* 371–375.

Thornton, M. C., Chatters, L. M., Taylor, R. J., & Allen, W. R. (1990). Sociodemographic and environmental correlates of racial socialization by Black parents. Child Development, 61(2), 401-409.

Trierweiler, S. J., Neighbors, H. W., Munday, C., et al. (2000). Clinician attributions associated with the diagnosis of schizophrenia in African American and non-African American patients. *Journal of Consulting and Clinical Psychology, 68,* 171–175.

Trierweiler, S. J., Muroff, J. R., Jackson, J. S., Neighbors, H.W., & Munday, C. (2005). Clinician race, situational attributions, and diagnoses of mood versus schizophrenia disorders. *Cultural Diversity & Ethnic Minority Psychology,11*(4), 351–364.

Trierweiler, S. J., Neighbors, H. W., Munday, C., Thompson, E. E., Jackson, J. S., Binion, V. J. (2006). Differences in patterns of symptom attribution in diagnosing schizophrenia between African American and non-African American clinicians. *American Journal of Orthopsychiatry, 76,* 154–160.

Turner, S. M., Hersen, M., & Heiser, N. (2003). The interviewing process. In M. Hersen & S. M. Turner (Eds.), *Diagnostic interviewing* (3rd ed.). New York: Wiley and Sons.

U.S. Department of Health and Human Services. (2001). Mental Health: Culture, Race, and Ethnicity-A Supplement to Mental Health: A Report of the Surgeon General. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

U.S. Census Bureau. (2010). About Poverty. Retrieved on January 3rd, 2014 from: http://www.census.gov/hhes/www/povrty/about/overview/index.html.

Vasquez-Barquero, J. L., Diez-Manrique, J. F., Pena, C., Aldama, J., Samaniego, R. C., Menendezm A. J., & Mirapeix, C. (1987). A community mental health survey in Cantabria: A general description of morbidity. *Psychological Medicine, 17,* 227–241.

Ventura, J., Liberman, R. P., Green, M. F., Shaner, A., & Mintz, J. (1998). Training and quality assurance with the Structured Clinical Interview for DSM-IV (SCID-I/P). *Psychiatry Research,79*(2), 163–173.

Vernon, S. W., & Roberts, R. E. (1982). Use of the SADS-RDC in a tri-ethnic community survey. *Archives of General Psychiatry,39*(1), 47.

Want, V., Parham, T. A., Baker, R. C., & Sherman, M. (2004). African American students' ratings of Caucasian and African American counselors varying in racial consciousness. Cultural Diversity and Ethnic Minority Psychology, 10, 123–136.

Ward, C. H., Beck, A. T., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1962). The psychiatric nomenclature. *Archives of General Psychiatry, 7,* 198–205.

Weathers, Frank W.; Ruscio, Ayelet Meron; Keane, Terence M. (1999). Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychological Assessment, 11*(2), 124–133.

Wells, K. B., Burnam, A., Leake, B., & Robins, L. N. (1988). Agreement between face to face and telephone administered versions of the depression section of the

NIMH Diagnostic Interview Schedule. *Journal of Psychiatric Research, 22,* 207–220.

Westmoreland, P., Gunter, T., Loveless, P., Allen, J., Sieleni, B., & Black, D. W. (2010). Attention deficit hyperactivity disorder in men and women newly committed to prison clinical characteristics, psychiatric comorbidity, and quality of life. *International Journal of Offender Therapy and Comparative Criminology,54*(3), 361–377.

Whaley, A. L. (1997). Ethnicity/race, paranoia, and psychiatric diagnoses: Clinician bias versus sociocultural differences. *Journal of Psychopathology & Behavioral Assessment, 19*, 1–20.

Whaley, A. L. (1998). Cross-cultural perspective on paranoia: A focus on the Black American experience. *Psychiatric Quarterly,69*(4), 325–343.

Whaley, A. L. (2001a). Cultural mistrust: An important psychological construct for diagnosis and treatment of African Americans. *Professional Psychology: Research and Practice,32*(6), 555.

Whaley, A. L. (2001b). Cultural mistrust and mental health services for African Americans: A review and meta-analysis. *The Counseling Psychologist, 29,* 513–531.

Widiger, T. A. (1997). Mental disorders as discrete clinical conditions: Dimensional versus categorical. In S. M. Turner & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (2nd ed., pp. 3–23). New York: Wiley.

Wilmink, F. W., & Snijders, T. A. B. (1989). Polytomous logistic regression analysis of the General Health Questionnaire and the Present State Examination. *Psychological Medicine, 19,* 755–764.

Williams, D. R., González, H., Neighbors, H., Nesse, R., Abelson, J. M., Sweetman, J., & Jackson, J. S. (2007). Prevalence and distribution of major depressive disorder in African Americans, Caribbean Blacks, and Non-Hispanic Whites: Results from the National Survey of American Life. *Archives of General Psychiatry, 64*(3), 305–315.

Wing, J. K., Birely, J. L. T., Cooper, J. E., Graham, P., & Isaacs, A. (1967). Reliability of a procedure for measuring and classifying present psychiatric state. *British Journal of Psychiatry*, *113*, 499–515.

Wittchen, Hans-Ulrich; (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research, 28*(1), 57–84.

Wing, J. (1996). SCAN and the PSE tradition. *Social Psychiatry and Psychiatric Epidemiology, 31*(2), 50–54.

Wing, J. K., Sartorius, N., & uSTUN, t. tb.(1998). Diagnosis and clinical measurement in psychiatry: A reference manual for SCAN/PSE-10. Cambridge UK: Cambridge University Press.

World Health Organization. (1994). *Schedules for Clinical Assessment of Neuropsychiatry (SCAN)*. Geneva: World Health Organization.

World Health Organization. (2004). The World Health Organization World Mental Health Composite International Diagnostic Interview (WHO WMH-CIDI). Retrieved on January 3rd, 2014 from: http://www.hcp.med.harvard.edu/wmhcidi/index.php.

Woodward, A. T., Taylor, R. J., Bullard, K. M., Neighbors, H. W., Chatters, L. M., Jackson, J. S., et al. (2008). The use of professional services and informal support among Black Americans with a mental disorder. *Psychiatric Services, 59*(11), 1292–1298.

Woodward, A. T., Taylor, R. J., Abelson, J. M., & Matusko, N. (2013). Major depressive disorder among older African Americans, Caribbean Blacks, and non-Hispanic Whites: Secondary analysis of the National Survey of American life. *Depression and Anxiety, 30*(6), 589–597.

Assessing Personality Using Self-Report Measures with African American Clients

4

Jessica Urgelles

The results of psychological assessments often inform important clinical decisions. Therefore, cultural bias in personality assessment is an important concern. Racial bias of assessment measures and in the interpretation of testing results continues to be a concern when using self-report personality measures with Blacks and African Americans. According to the U.S. Census (2013), Black Americans make up 13 % of the US population. The history of Africans in the USA of American includes over 200 years of slavery, the effects of which still have negative consequences for many African Americans today. Twenty-eight percent of Black Americans currently live in poverty, compared to only 9.8 % of non-Hispanic Whites (Denavas-walt et al. 2012). In 2011, Black Americans earned the lowest income of any other group when compared to other races and those of Hispanic origin. The unemployment rate for Black Americans (16 %) is higher than it is for other groups (U.S. Census Bureau 2012a) and a disparate number of Black Americans do not attain higher education (U.S. Census Bureau, 2012b). Black Americans are also incarcerated at disproportionate rates (U.S. Census Bureau 2011a). These data are important for clinicians to consider because sociodemographic factors have been statistically linked to psychopathology (Aranda et al. 2012; Barnes 2008; Gibbs et al. 2013; Regier et al. 1993). Indeed, socioeconomic status, Black race, and marital status are among the biggest risk factors for developing a personality disorder (Grant et al. 2004).

Barnes (2008) found that African Americans admitted to a state psychiatric hospital were four times more likely to get a schizophrenia diagnosis than Caucasians, even after controlling for demographic variables. On the other hand, African Americans were less likely to receive a diagnosis of major depressive disorder and bipolar disorder. Barnes (2008) attributes these discrepancies to a lack of cultural competence in diagnosing ethnic minorities and not actual differences in prevalence rates. This is consistent with data from the Epidemiologic Catchment Area studies which demonstrated differential relationships between psychiatric disorders and various demographic characteristics, such as age, gender, marital status, race or ethnicity, and socioeconomic status (Regier et al. 1993). The authors found that the prevalence of any disorder was higher for African Americans than it was for any other race or ethnic group. However, when the odds ratios were adjusted for other demographic variables, the prevalence for psychiatric disorder in African Americans was no different than in other groups, with the exception of severe cognitive impairment. The authors speculated that there were differences in educational experiences between the ethnic groups that were not well captured by their socioeconomic variable. The demographic variables

---

J. Urgelles (✉)
Department of Psychology, University of Nevada,
Las Vegas 4505 S, Maryland Pkwy, Box 455030,
Las Vegas, NV 89154-5030, USA
e-mail: jurgelles@msn.com

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_4,

most associated with psychiatric disorder were socioeconomic and marital status. In addition, there is evidence that differences in sociodemographic variables are associated with differences in personality traits (Al-Halabí et al. 2010). Based on these findings, there has been concern that personality measures could overpathologize African Americans and other ethnic minorities because of cultural differences, but also because they tend to be overrepresented in low-socioeconomic status groups (Hall et al. 1999).

Cultural Considerations for Practitioners

Blacks in the USA may be immigrants from Africa or the Caribbean. These families may not have experienced slavery like African Americans have, but they may have similar experiences with racism and oppression. Mental health differences between African Americans and Caribbean Blacks are rarely studied. Instead, they are often grouped together in psychology research. The National Epidemiological Survey of Alcohol and Related Conditions is one of the only studies to investigate the prevalence rates of mental health disorders separately for African Americans and Caribbean Blacks (Gibbs et al. 2013). In this study, African Americans and Caribbean Blacks had equal or lower lifetime prevalence rates of Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV) Axis I disorders, but higher prevalence rates of personality disorders, compared to non-Hispanic Whites. When African Americans and Caribbean Blacks were compared, Caribbean Blacks had a higher 12-month prevalence of psychotic disorders and a lower lifetime prevalence of major depressive disorder, alcohol dependence, and drug abuse than African Americans. Although this study highlights the importance of considering African American and Caribbean Blacks as separate cultural groups, currently there is no published data that does so in the personality assessment literature. Therefore, the research presented in this chapter does not differentiate between the two groups. Nevertheless, clinicians should be sensitive to potential differences between African American and Caribbean Blacks in their clinical practice.

True for both African Americans and Caribbean Blacks, an individual's unique racial experiences will shape their world view, including their self-identity and behavior (Boyd-Franklin and Karger 2012). Additionally, with the increase in biracial marriage, an individual's racial identity is much more complex than the color of their skin. Racial beliefs and experiences can be passed down through families in order to instill racial pride, combat negative images, and prepare children for racism and discrimination, which are part of normal racial socialization (Boyd-Franklin and Karger 2012). In addition, beliefs about racial experiences can change throughout an individual's lifespan through the process of acculturation and reshaping their worldview and beliefs about self and others (Cross 1978).

Cross (1978) proposed a five stage model of normative Black acculturation: (a) Preencounter, (b) Encounter, (c) Immersion-Emersion, (d) Internalization, and (e) Internalization-Commitment. Depending on where an individual is on this progression, they could experience self-hatred, glorification of Black culture and destructiveness towards the dominant culture, flexibility, openness, or self-confidence. Therefore, there could be factors that can affect personality assessment profiles that are influenced by normal racial socialization or the state of an individual's acculturation status, potentially pathologizing a normative process for that individual.

In additional to racial socialization and acculturation, the effects of racial discrimination should also be considered in the personality assessment of African Americans. Racial discrimination has been shown to contribute a significant proportion of the variance in Blacks' psychiatric symptoms (Klonoff et al. 1999). However, clinicians must be careful not to interpret the functioning of African Americans as solely a response to racism (Harrell 2000). A combination of racial discrimination, oppression, and poverty in the lives of African Americans has resulted in anger and rage among some African Americans (Boyd-Franklin and Karger 2012). Some researchers have interpreted this to mean that there is a higher incidence of psychopathic personality among African Americans compared to other racial and ethnic groups (Lynn 2002). However, others argue that antisocial behavior is less related to personality in Blacks, and more associated with other variables such as socioeconomic status and historical circumstance (Skeem et al. 2003; Zuckerman 2003). Practitioners are urged to consider the potential effects of racial socialization, acculturation, and racial discrimination on personality profiles and not assume personality profile scores necessarily reflect stable traits.

Another important consideration for practitioners is the validity of assessment measures. Using a personality assessment measure with African Americans assumes that the same traits can be used to describe personality in African American culture and the culture in which it was developed (McCrae and Terracciano 2006). In addition, a given score has to be similarly meaningful for African Americans as it is for the other cultures. One way of determining the validity of a personality measure for use with a specific culture is to compare it to other measures that are validated for use with that culture (McCrae et al. 2005c). However, there are no personality measures developed and validated specifically for use with African Americans so clinicians will have to make their own informed decisions regarding which measures to use. The following sections will review the literature and make recommendations regarding the use of self-report personality instruments with African Americans. The focus will be on the most widely used self-report personality measures, including the Minnesota Multiphasic Personality Inventory-2 (MMPI-2), Millon Clinical Multiaxial Inventory-III (MCMI-II), Personality Assessment Inventory (PAI), and NEO.

The MMPI-2

The MMPI-2 (Butcher et al. 2001; Butcher et al. 1989) is one of the most widely studied and utilized self-report personality measures (Hall et al. 1999; Hall and Phung 2001; Kwan and Maestas 2008). The MMPI-2 is a 567 true–false test, recommended for use with adults (18 years or older) who have at least a sixth grade reading level. The MMPI-2 is made up of many scales and subscales, some of which are better validated than others. It has several validity scales that assess inconsistent (cannot say, variable response inconsistency, true response inconsistency) and biased responding (infrequency, lie), as well as a scale that provides a correction factor for the over- or underreporting of symptoms (correction). There are also ten clinical scales that generally align themselves with Diagnostic and Statistical Manual of Mental Disorders (DSM) syndrome criteria. There are also a great number of supplemental and content scales that have been developed and normed by the publishers. There is a wealth of data on the Minnesota Multiphasic Personality Inventory (MMPI) from many years of research (e.g., Arbisi et al. 2002; Ben-Porath et al. 1995; Castro et al. 2008; Dana and Whatley 1991; Davis 1975; Frueh et al. 1996; Greene 1987; Hall et al. 1999; Harrison and Kass 1967; McNulty et al. 1997; Timbrook and Graham 1994; Walters et al. 1983; Whatley et al. 2003).

The original MMPI did not include African Americans in the normative sample (Graham 2006), and thus one of the goals during the development of the MMPI-2 was to obtain a large normative group that was representative of the national population to include African Americans (Butcher et al. 1989). The makeup of the normative sample was guided by the 1980 Census data, resulting in 12% of the sample being African American. Although this sample was not matched on demographic characteristics, summary data for each of the ethnic groups included in the normative sample is presented in Appendix H of the MMPI-2 manual (Butcher et al. 1989). When one looks at the normative data it is evident that African Americans scored somewhat higher than Caucasians on most clinical scales, but that only scale 4 (Psychopatic/Deviate) for women demonstrated a meaningful difference (Butcher et al. 1989; Greene 1987). Given that differences between African Americans and Caucasians on the MMPI were almost nonexistent when demographic variables were taken into account, the MMPI restandardization committee concluded that separate norms were not needed for African Americans (Graham 2006).

In a sample of 106 African American and 131 Caucasian undergraduate students completing both the MMPI and the MMPI-2, African Americans scored significantly higher than Caucasians on both versions of the measure on scales 6, 8, and 9 (Goldman et al. 1995). The racial differences were reduced, particularly on the MMPI, after controlling for socioeconomic status, but remained statistically significant on scales 6 and 8. Mean T scores were generally lower for both groups on the MMPI-2 compared to the MMPI, with mean scores for African American males showing the greatest reduction. The authors concluded that the more diverse normative sample of the MMPI-2 reduces mean score differences attributable to race, but that the effects of racial socialization and alienation by the dominant culture are still evident.

A meta-analysis by Hall et al. (1999) examined the results of 31 years of MMPI and MMPI-2 research. It included 25 studies of African American and Caucasian men and 12 studies of African American and Caucasian women. They found that African American men scored higher than Caucasian men on the L, F, and K validity scales, and clinical scales 1, 7, 8, and 9. Caucasian men scored higher on clinical scales 2, 3, 4, 5, and 0. African American women scored higher than Caucasian women on the L and F validity scales and clinical scales 1, 2, 4, 5, 6, 7, and 8. They scored lower than Caucasian women on scales K, 2, 3, and 9. The authors determined that the effect sizes associated with all of these differences were small and not clinically meaningful. None of the mean scale differences were greater than 5 T-score points.

MMPI Studies with the Normative Sample

The MMPI-2 normative sample was reexamined for bias against those of lower-socioeconomic status. Long et al. (1994) found that differences in MMPI-2 scores associated with income and education were small (less than 5 T-score points) and therefore not clinically meaningful. There were differences between scores on some of the validity scales that were equal to or greater than five points. Those is the low-socioeconomic status levels tended to score higher on the F scale and lower on the K scale than those in the higher-socioeconomic status levels. Despite these differences, conclusions regarding test validity did not differ. An examination of extra-test characteristics revealed that the MMPI-2 tended to underpredict symptoms for lower-socioeconomic status participants and overpredict symptoms for higher-socioeconomic status levels. This directly contrasts the concern that the MMPI-2 would overpathologize lower-socioeconomic status individuals and therefore African Americans.

A subset of the MMPI-2 standardization sample and a clinical inpatient group were examined to determine the relationship between various demographic factors and MMPI-2 scores (Schinka et al. 1998). The authors found that demographic variables contributed little incremental variance beyond that explained by the presence of psychopathology for both the validity and clinical scales. Of all the demographic variables they examined, gender explained the most variance over other variables, including ethnic identification.

Timbrook and Graham (1994) also reexamined a subset of the MMPI-2 normative sample to look for evidence of bias based on race. They found that African American men scored significantly higher than Caucasian men on scale 8, even after the sample was matched for age, years of education, and total family income. When women were compared, the means for scales 4, 5, and 9 were significantly higher for African Americans than they were for Caucasians. However, the mean difference of all scores was less than 5 T-score points, suggesting that these differences were not clinically meaningful. The authors also investigated the validity of MMPI-2 scores. They calculated how much MMPI-2 scores predicted conceptually related extra-test information provided by participants' partners for five of the clinical scales. Although not statistically significant, the data demonstrated that the MMPI-2 tended to underpredict symptoms and problems of African Americans in the normative sample, the opposite of what was expected. The only statistically significant difference in predictive accuracy was scale 7, which tended to underpredict ratings of anxiety for African American women. Based on these findings, the authors concluded that the MMPI-2 did not appear to be biased against African Americans.

MMPI Studies with Psychiatric Outpatient Samples

The validity of the MMPI-2 was examined in a sample of African American and Caucasian clients in a community mental health center (McNulty et al. 1997). Of the 684 participants in the sample, 123 were African American. African American men scored significantly higher than Caucasian men on the L scale and the FRS (fears) content scale and African American women scored significantly higher on scale 9 and the LSE (low self-esteem) content scale than Caucasian women. MMPI-2 scores were also compared with conceptually-related therapist ratings. No significant differences between racial groups were found in correlations between MMPI-2 scores and therapist ratings. The authors concluded that although there were some mean scale score differences between the racial groups, the MMPI-2 was not biased against African Americans in their outpatient mental health setting because of the similar relationships between extra-test variables and MMPI-2 scores for the two racial groups. Similar findings have been demonstrated in the only study examining the use of the Restructured Clinical Scales with African Americans (Castro et al. 2008).

The difference between the MMPI-2 scores of 88 African American and 118 Caucasian male combat veterans at a posttraumatic stress disorder (PTSD) outpatient treatment program were examined and it was demonstrated that African Americans scored higher than Caucasians on the F-K Index and scales 6 and 8 (Frueh et al. 1996). The two groups did not differ on separate measures of depression and PTSD symptoms, but African Americans scored higher on a measure of dissociation. The F-K index is a validity indicator that detects faking bad profiles, though it is not as effective as the F scale (Graham 2006). Other measures of paranoia and psychotic symptoms were not used so it could not be determined if African Americans actually had more severe psychopathology than Caucasians. The authors hypothesize that the African American veterans in this sample may have exaggerated their symptoms in order to get help from a predominantly Caucasian hospital staff. A follow-up study conducted by Frueh et al. (1997) evidenced no differences between African American and Caucasian combat veterans with a diagnosis of PTSD on any of the MMPI-2 scales. The two groups also did not differ on any of the additional measures of psychopathology.

Dean et al. (2008) examined scores on the Faking Bad Scale (FBS) in a sample of 190 psychiatric outpatients (23 % African American). They were concerned that non-White ethnic groups could elevate the FBS due to a tendency to report more somatic symptoms. However, there were no significant racial or ethnic differences on FBS performance in their outpatient sample.

MMPI Studies with Psychiatric Inpatients

Some researchers have concluded that the MMPI-2 is not biased against African Americans in psychiatric inpatient samples (Arbisi et al. 2002; Stukenberg et al. 2000). Aribisi et al. (2002) examined the MMPI-2 scores of 229 African American and 1558 Caucasian inpatients at a veteran's affairs hospital and an urban medical center. They found that African Americans scored higher than Caucasians on one validity scale and several clinical, content, and supplemental scales, including Scales F, 4, 6, 8, 9, FRS, Depression (DEP), Health Concerns (HEA), Bizarre Mentation (BIZ), Anger (ANG), Cynicism (CYN), Antisocial Practices (ASP), Family Problems (FAM), Negative Treatment Indicators (TRT), MacAndrew Alcoholism Scale—Revised (MAC–R), and Addiction Acknowledgement Scale (AAS). Caucasians scored higher on scale K. With the exception of scale K, all differences were greater than 5 T-score points and therefore clinically significant. However, these differences were generally consistent with extra-test characteristics. The authors performed step-down hierarchical multiple regression analyses predicting conceptually relevant clinical criteria to test for potential test bias. Bias was found in several scales, including scales 2, 4, 8, 9, DEP, ASP, AAS, APS, and MAC–R; however, the effect sizes were at most small. Additionally, the bias was found in the opposite direction, with a slight underprediction of psychopathology in African Americans. The authors argue that the MMPI-2 scales can still be interpreted the same way for both groups because (1) the MMPI-2 predictor scale still accounted for most of the variance in prediction for all of the criterion variables, (2) the mean differences between the racial groups on MMPI-2 scales were also apparent in the extra-test variables, and (3) the effect size of race on prediction was consistently small.

More recently, a study with 1411 veteran psychiatric inpatients (61 % African American) seeking help for substance abuse problems showed evidence of a prediction bias such that the MMPI-2 tended to overpredict psychopathology in African Americans (Monnot et al. 2009). However, the effect sizes were small, accounting for 3–5 % of the variance. The authors caution that the MMPI-2 may be a biased personality assessment measure when working with certain populations (i.e., male veteran substance abusers).

Similar to previous studies, Munley et al. (2001) found that African American men scored higher than Caucasian men on the FRS, CYN, ASP, and BIZ scales in a veteran inpatient population. This sample consisted of 180 African American and 180 Caucasian matched pairs. The authors speculated that difference on the BIZ scale between the two races could be due to differences in cultural beliefs about visions, ghosts, and spirits. This finding was replicated in another study with veteran inpatients (Monnier et al. 2002). However, African Americans in the latter study were more likely than Caucasians to receive a psychotic disorder diagnosis based on a clinical interview.

MMPI Studies in Forensic Samples

In a sample of 137 Caucasian and 47 African American men undergoing court-ordered forensic psychological evaluations, the mean scores on the MMPI-2 validity, clinical, and content, and substance abuse scales were remarkably similar for both African Americans and Caucasians (Ben-Porath et al. 1995). The only significant differences were found on the CYN and ASP content scales, with African Americans scores higher than Caucasians. Follow-up analyzes of individual items revealed that African Americans tended to endorse more items that expressed skepticism of the good motives of others. They also endorsed more items suggestive of antisocial attitudes, such as disregard for the law. The two groups did not differ in their report of specific antisocial behaviors. Extra-test data were not available to examine whether or not the African American men in this sample actually have more cynical and antisocial attitudes than Caucasians. The groups were not matched for demographic characteristics, but the socioeconomic statuses of the groups were not statistically different. These results could be attributed to the effects of racism, resulting in more distrust of others.

A study with 217 male pretrial defendants (46 % African American) also demonstrated higher scores for African Americans on the CYN content scale compared to Caucasians (Shea et al. 1996). African American participants in this study also scored higher than Caucasians on Scale 9. In addition, African Americans scored lower than Caucasians on clinical scales 1, 2, 3, and 0, and content scales ANX, DEP, HEA, LSE, and social discomfort (SOD).

A forensic study with 51,486 inmates (55 % African American) found that American Americans had a significantly greater number of invalid profiles compared to Caucasians (McNulty et al. 2003). In particular, African Americans were more likely respond to items indiscriminately, by either responding randomly or generally responding true or false, which resulted in higher CNS and VRIN scores than Caucasians. This type of response style is associated with lower reading skills, attention problems, or uncooperativeness. There were no differences between the two racial groups related to education level in this study, so the authors speculate that the discrepancy in response style between African Americans and Caucasians was associated with a difference in motivation.

African Americans in a maximum security forensic inpatient hospital scored significantly higher than Caucasian forensic inpatients on the Overcontrolled-Hostility (O-H) scale (Hutton et al. 1992). The sample consisted of 136 African Americans and 276 Caucasians. The authors warned that use of the O-H scale with African Americans could lead to a bias of labeling them as aggressive or violent. However, the O-H scale failed to correlate with behavioral indicators of overcontrolled-hostility for both racial groups, suggesting that it has questionable validity.

MMPI-2 Substance Abuse Studies

There is evidence that the MacAndrews alcoholism (MAC) scale should not be used with African Americans because it fails to distinguish African American alcoholics from African American nonalcoholics (Walters et al. 1984, 1983). Studies on the MAC scale have found that nonalcoholic African American men tend to score relatively high MAC scores, making it difficult to distinguish them from alcoholics (Walters et al. 1984, 1983). However, studies on the MAC scale have used military populations, so it is unclear if the results generalize to other populations. The MAC also has not been well studied in female populations (Graham 2006). In a sample of outpatients, the MAC had little practical use for classifying Caucasian, African American, male, female substance abusers (Gripshover and Dacey 1994). The revised version of the MAC scale, the MAC-R, only varies by four questions and has not been well tested with African American populations. However, one study with adolescent inpatients demonstrated that accuracy of classification was similar for African Americans and Caucasians (Micucci 2002). Given the available research, extreme caution should be used when interpreting high scores on the MAC-R with African American clients and efforts should be made to obtain corroborating information.

MMPI-2 and Acculturation

Utilizing unpublished MMPI data, it was demonstrated that the acculturation status of 50 African American male college students predicted MMPI scale scores (Whatley et al. 2003). Acculturation was measured using the Racial Identity Attitude Scale—Black, Short Form (RIAS–B). Immersion–Emersion scale scores of the RIAS–B predicted scores on scales 4 and 9. A 4–9 code type is frequently associated with antisocial personality attributes. This suggests that some of the variance in this code type could be associated with a normative racial identity pattern in which the individual adopts a more Afrocentric orientation. The RIAS-B internalization scale predicted MMPI scores on scale 6. This suggests that internalizing African American culture and values, while maintaining openness and flexibility, could be associated with less distrust of others. Studies of this nature have not been conducted with the MMPI-2, but this study demonstrates the need for more research that investigates the relationship between the MMPI-2 and normal psychological processes associated with African American identity.

Minnesota Multiphasic Personality Inventory-Adolescent (MMPI-A)

There is a dearth of cultural research on the MMPI-A, the adolescent version of the MMPI-2. One study found that the MMPI-A was about equally effective at diagnosing substance abuse in 39 African American, 35 Caucasian, 3 Hispanic, and 2 mixed race psychiatric inpatient adolescents (Micucci 2002). In a forensic sample of 54 adolescent first-time offenders, African Americans scored higher on the repression scale than Mexican Americans (Gómez et al. 2000). African Americans also had a greater percentage (50 %) of within-normal-limits profiles than Mexican American (25 %). At this point, there is not enough research available on the MMPI-A to make recommendations regarding its use with

African Americans. It should be used with caution in this population and should not be the only measure utilized for diagnostic purposes.

Summary

The MMPI-2 is a widely used self-report measure of personality that is well researched in many different populations, including psychiatric inpatients and outpatients, veterans, forensic populations, substance users, and nonclinical samples. There is evidence that African Americans score higher that Caucasians on most scales, particularly scales 6, 8, 9, FRS, CYN, ASP, and BIZ. However, these differences were typically small and nonsignificant after controlling for demographics and diagnoses. Educational differences may be associated with African Americans receiving higher scores on some of the validity scales. Several studies showed that MAC-R may not be a valid measure of substance abuse for African Americans; therefore, it is not recommended for use with this population. Researchers generally concluded that the MMPI-2 is a valid measure of personality for African Americans, but there is evidence that it may underpredict pathology in African Americans. There is an adolescent version of the MMPI-2, the MMPI-A, which has not been well researched with African American samples. Based on the available research, the MMPI-2 appears to be a valid measure of personality for assessing African Americans, but practitioners should be aware that scores may be an under prediction of true levels of psychopathology. The use of additional measures and sources of information are recommended to corroborate findings on the MMPI-2.

MCMI-III

The MCMI-III (Millon et al. 2006, 2009; Millon 1994) was developed for use with clinical populations and should not be used with individuals who do not have mental health problems. It was developed based on a complex theory of personality development and has been updated over the years to reflect advancements in personality theories, psychopathology research, and clinical data (Millon et al. 2009). The result is a measure consisting of 175 items that make up 11 scales that assess clinical personality patterns (Schizoid, Avoidant, Depressive, Dependent, Histrionic, Narcissistic, Antisocial, Aggressive, Compulsive, Negativistic, Masochistic), 3 scales that assess severe personality pathology (schizotypal, borderline, paranoid), 7 scales that assess clinical syndromes (Anxiety Disorder, Somatoform Disorder, Bipolar: Manic Disorder, Dysthymic Disorder, Alcohol Dependence, Drug Dependence, Posttraumatic Stress Disorder), 3 scales that assess severe clinical syndromes (Thought Disorder, Major Depression, Delusional Disorder), and 5 modifying indices that assess validity (Disclosure, Desirability, Debasement, Invalidity, Inconsistency). Unique to the MCMI is a sophisticated method of accounting for individual scale elevations that accounts for the interrelationships among the scales. This is accomplished via the Grossman Facet Scales that clarify elevations by identifying personality processes and domains that contribute to scale elevations, and a base rate (BR), rather than a T score, which is derived for each scale based on the prevalence of characteristics and patterns in the psychiatric population for a given personality disorder or syndrome. Scores range from 0 to 115 (Millon et al. 2009). For the personality disorder scales, a BR of 75 indicates the presence of a trait and a score of 85 indicates the presence of a disorder. A score of 75 on the clinical syndrome scales indicates the presence of a syndrome and a score of 85 indicates that a syndrome is prominent. Several changes were made to the MCMI with the publication of the MCMI-III, including the addition of two scales, changing 95 of the 175 items, and changing the items responses from a 3-point scale to a 2-point scale (Millon 1994). Despite these changes, it has been demonstrated that the MCMI-III has a similar factor structure to the MCMI and the MCMI-II, even with African American samples (Craig and Bivens 1998).

Research on the MCMI and its utility and validity with non-Caucasian groups has been conspicuously absent since the 1990s, with only one study conducted since the publication of the MCMI-III. African Americans were included in the normative sample of the MCMI-III, but they were underrepresented in both the scale development (8.7 %, N = 52) and cross validation (8.3 %, N = 33) samples (Millon et al. 2006). The MCMI-III was recently renormed with a sample of 83 African Americans (11.1 %) which is still not representative of the national population (Millon et al. 2009).

MCMI Studies with Black Norms

Separate norms for African Americans and other non-Caucasian samples were developed for the original MCMI, but they were based on small samples (Millon 1984). When the MCMI-II was published in 1987 it only contained norms based on sex, even though there had been no relevant published research that might have led to this modification (Davis et al. 1990).

The separate Black and Caucasian norms were later evaluated to determine their ability to distinguish male patients with a diagnosis of schizophrenia from those without a diagnosis of schizophrenia at a veteran's hospital (Davis et al. 1990). It was determined that when the Black norms were used African Americans were identified as significantly more pathological that Caucasians, even after controlling for education. African Americans scored significantly higher on four out of the five scales examined, including the antisocial, avoidant, psychotic thinking, and Psychotic Delusions Scales.

Choca et al. (1990) had similar results using the Black norms with a sample of Caucasian and African American males who were psychiatric inpatients at a Veteran's Affairs medical center. The participants were 235 African American and 471 Caucasian males. The authors found that the MCMI was more likely to diagnose African Americans with all disorders, except for personality disorders, and that African Americans had significantly different scores on 9 of 20 scales. African Americans scored significantly higher on the histrionic, narcissistic, antisocial, paraphrenia, hypomania, alcohol abuse, drug abuse, and psychotic delusion scales. Caucasians scored significantly higher on the dysthymia scale. Possible bias was also found at the item level. When 209 Caucasian and 209 African American participants were matched according to psychiatric diagnosis, it was demonstrated that the two racial groups responded differently to 45 of the 175 items. However, a principal components factor analysis revealed that the factor structures for the two racial groups were identical. This led the authors to support the use of the MCMI with African Americans. They recommended that the Black norms be adjusted and hypothesized that the Black norms in the MCMI test manual (Millon 1984) did not adequately reflect the racial differences in the overall population. The authors also noted that due to few changes between the MCMI and the MCMI-II, the concerns regarding racial bias are probably similar for both measures.

It was later determined that the Black norms were useful for distinguishing male psychiatric inpatients from a veteran's hospital who self-identified as angry or having psychotic symptomatology (Greenblatt and Davis 1992). The participants were 778 Caucasian and 272 male psychiatric inpatients. The authors found that the MCMI was equally accurate in predicting angry and psychotic patients for both races utilizing the Black norms for African American participants, with both races exhibiting very similar scale elevations. The passive-aggressive scale best predicted anger and the avoidant scale best predicted psychosis for both races.

MCMI Studies with Combined Norms

Few studies have examined the use of the MCMI combined norms with African Americans, with only one study published since the 1990s. Studies on the older versions of the MCMI demonstrate that it may be biased against African Americans. It was found that the Drug Abuse scale was more likely to identify African American participants, narcissistic/antisocial subtypes, and those with more severe psychopathology, as drug abusers (Calsyn et al. 1991). The authors cautioned against the use of this scale. Similarly, Hamberger and Hastings (1992) showed that even after controlling for age, education, and employment status, African Americans scored higher than Caucasians on the Narcissistic, Aggressive, Paranoid, Drug, and Psychotic Delusion Scales, and warn that interpreting the MCMI scores of African Americans without considering the cultural context could result in an over reporting of pathology for African Americans. They point out that experiences with prejudice and oppression and differential socialization could help explain differences between the personality test scores of African Americans and Caucasians. Additional research has showed that African Americans score higher than Caucasians on the Histrionic, Narcissistic, Paranoid, Drug Dependent, And Delusional Disorder scales of the MCMI-II (Munley et al. 1998). However, these differences became nonsignificant after a smaller subset of the racial groups were matched on Axis I discharge diagnosis and substance abuse comorbidity. The authors recommended that participants be matched for both Axis I and Axis II diagnoses in future research. These studies demonstrate that the older versions of the MCMI, from which the current versions were developed, have had evidence of possibly overpathologizing African Americans.

In the only publication to examine the use of the MCMI-III with diverse groups, African American combat veterans did not have significantly different MCMI-III personality profiles than Caucasian combat veterans (Ghafoori and Hierholzer 2010). However, these findings were based on a very small sample of African Americans ($n = 12$).

The Millon Adolescent Clinical Inventory (MACI)

The MACI (Millon 1993) is the adolescent version of the MCMI-III and is intended for use with adolescents between the ages of 13 and 19. Unfortunately, it has not been well researched with diverse populations. One study examined the MACI profiles of 103 adolescent males (60.2 % African American) at a youth detention center (Stefurak et al. 2004). The authors found that there was no relationship between personality clusters determined by the MACI and participant race. The Millon Pre-Adolescent Clinical Inventory and the Millon Adolescent Personality Inventory are also available, but neither has been evaluated with diverse populations.

Summary

The MCMI-III is a self-report personality measure developed for use with clinical populations. African Americans were not well represented in the development and standardization samples of this measure. Black norms were developed for the MCMI, but were not published for newer versions of the measure. In addition, its use with non-Caucasian samples has not been researched in recent years. Only one study has been published since the 1990s and the publication of the MCMI-III, but this study included a very small number of African Americans in its sample. Research with the MCMI and MCMI-II has demonstrated that African Americans tend to score higher on several scales, particularly the Narcissistic, Aggressive, Paranoia, Drug Abuse, And Psychotic Delusion scales. Despite these differences, the two racial groups appear to have the same factor structure. Concern has been raised about the validity of the Drug Abuse Scale in African American samples, and the use of this scale with African Americans is not recommended. The adolescent version, the MACI, has not been well studied in nonCaucasian samples. Given these findings practitioners are cautioned against using the MCMI-III with African American populations until it can gain more support from future research.

PAI

The PAI (Morey 1991, 2007a) was developed to be a psychometrically sound and culturally unbiased measure of personality. To accomplish this goal a bias review panel was assembled with professionals and citizens from varying backgrounds, including Black male and female psychologists and Black male and female citizens

(Morey 1991). Any potentially biased, offensive, or confusing items were removed or revised. The resulting measure consists of 344 items with 11 clinical scales (Somatic Complaints, Anxiety, Anxiety-Related Disorders, Depression, Mania, Paranoia, Schizophrenia, Borderline Features, Antisocial Features, Alcohol Problems, and Drug Problems), 5 treatment consideration scales (Aggression, Suicidal Ideation, Nonsupport, Stress, and Treatment Rejection), 2 interpersonal scales (Dominance and Warmth), and 4 validity scales (Infrequency, Inconsistency, Negative Impression Management, And Positive Impression Management). The PAI was developed for use adults over the age of 17. The PAI item booklet is written at a fourth-grade reading leveling, allowing for its use with adults with a wide range of educational levels.

The normative sample contained individuals from both community and clinical settings. The community sample consisted of 1000 census-matched individuals, selected on the basis of cross-stratification for the variables of gender, race, and age (Morey 1991). The U.S. Census projections for adults 18 years and older for the year 1995 were used, which resulted in 117 (11.7 %) Black participants in the community sample. This is slightly below more recent estimates of the US adult Black population, at about 12.5 % (U.S. Census Bureau 2011b). The clinical sample contained 1265 patients, 12.6 % of which were Black or African American (Morey 1991).

In the normative sample, Black subjects tended to score higher than White participants on most scales, but these differences were quite small, in most cases equivalent or less than the standard error of measurement (Morey 1991). The Paranoia scale was the one notable exception. Black and other non-White participants in the normative sample scored approximately seven T score points higher than White participants. Morey (1991) attributed this finding to the prejudice experienced by non-White populations, which could lead to heightened vigilance and feelings of being treated unfairly. The Appendix of the PAI manual (Morey 1991, 2007a) contains separate norms based on the data from the census-matched Black community sample. Morey (1991) noted that the separate Black norms can be useful for the purposes of making comparisons, but that the T scores derived from the full normative data are strongly recommended for clinical and research purposes because they are based on a larger and more representative sample.

PAI Studies

The reliability and validity of the PAI was examined with a sample of lower-socioeconomic, methadone maintenance patients (Alterman et al. 1995). The sample was 49 % Hispanic, 46 % African American, and 5 % Caucasian. In this study sample, 30 % of the PAI profiles were invalid (Alterman et al. 1995), which is markedly higher than the 12–15 % rate reported for the clinical standardization sample (Morey 1991). This raises questions about the appropriateness of score cut-offs to determine invalid profiles in this population. The invalid profiles from this study were examined in a later study and it was found that those with invalid profiles on the PAI had similar patterns of response distortion on two structured interviews, the Addiction Severity Index and the Diagnostic Interview Schedule (Alterman et al. 1996). This suggests that invalid profiles were not obtained due to difficulty interpreting items on the PAI.

The psychometric properties of the PAI for this study sample were similar to those for the normative sample (Alterman et al. 1995). However, the internal consistency of one clinical scale (Drug Problems), two treatment consideration scales (Nonsupport and Treatment Rejection), and one interpersonal scale (Warmth) were questionable in this sample. The scale intercorrelations were similar to those demonstrated by the normative sample. Evidence of concurrent validity was obtained with comparisons of the PAI data with data from the Diagnostic Interview Schedule and the Addictions Severity Index.

The PAI scores of the study sample were compared with those from three other samples, the normative sample, the clinical standardization sample, and a sample of 158 substance dependent patients whose scores were provided to the investigators by the PAI developer via a personal communication (Alterman et al. 1995). The study sample had significantly higher scores than the normative sample on all clinical, treatment consideration, and interpersonal scales except for the Treatment Rejection and Warmth scales, in which they scored significantly lower. When the study sample was compared to the clinical standardization sample it was found that the study sample had significantly higher scores on the Mania, Paranoia, Antisocial Features, Drug Problems, Aggression, Dominance, and Warmth Scales, while the clinical standardization sample had significantly higher scores on the Anxiety, Depression, Alcohol Problems, and Suicidal Ideation Scales. The study sample scored significantly higher than the substance dependent sample on the Dominance scale, and significantly lower than the substance dependent sample on the Alcohol Problems and Suicidal Ideation scales. The scores were otherwise similar to those of the substance dependent sample. The profiles of the study sample were compared to the profiles of each of the three other samples. There was essentially no relationship with the clinical scale profiles of the normative sample (0.03), a moderate relationship with the clinical sample (0.45), and a strong relationship with the substance dependent sample (0.81). The coefficients for the treatment consideration scales were 0.28, 0.85, and 0.92 when study sample was compared with the normative, clinical, and drug dependent samples, respectively. This demonstrates that the PAI scores of this diverse sample were similar to those of other clinical and substance dependent samples.

The convergent and discriminant validity of the PAI Alcohol Problems and Drug Problems Scales were examined in a sample of 103 male veterans in a residential alcohol and drug treatment facility (Parker et al. 1999). The sample was 67 % African American and 33 % Caucasian. Internal consistency was good for both the Alcohol Problems (0.92) and the Drug Problems Scales (0.78), though the alpha for the Drug Problems Scale was somewhat lower than what was reported for the clinical standardization sample. Convergent validity was demonstrated by significant correlations between the PAI Drug Problems Scale and the Addiction Severity Index Drug Composite score (0.39), and the PAI Alcohol Problems and the Addiction Severity Index Alcohol Composite score (0.49). Convergent validity was also evident in point biserial correlations between the PAI Drug Problems Scale and drug-related discharge diagnoses (0.47) and between the PAI Alcohol Problems Scale and alcohol-related discharge diagnosis (0.47). An examination of the correlations between the Addiction Severity Index Composite Scales and the PAI clinical scales revealed evidence of discriminant validity of the Drug Problems and Alcohol Problems Scales. Therefore, this study demonstrated that the PAI Drug Problems and Alcohol Problems Scales had good psychometric properties in a predominantly African American clinical sample.

The PAI was more recently used in a study with male and female jail inmates (Youman et al. 2010). This sample was 55 % African American and 45 % Caucasian. In this study, African American and Caucasian inmates were equally in need of mental health treatment, defined as an elevation on at least one clinical scale of the PAI. However, there was a difference in the types of mental illness symptoms reported by the two races. African Americans inmates tended to endorse symptoms of persecution, grandiosity, and egocentricity. Caucasians were prone to affective problems, substance use, and impulsivity. The authors suggest that these test differences appear to be due to the effects of racism in African Americans and the development of protective factors, rather than evidence of test bias.

Personality Assessment Inventory-Adolescent (PAI-A)

The PAI-A (Morey 2007b) is the adolescent version of the PAI and is designed for use with adolescents between the ages of 12 and 18 years. The community standardization sample was comprised of 707 adolescents and fully crossed for age, gender, and race/ethnicity based on the 2003 US census. African Americans made up 15.4 % of the community sample. The clinical sample consisted of 1160 adolescents, 19.8 % of which were African American. Norms are available for the community and clinical samples. Separate norms are not available for African Americans. Just like the standardization sample of the PAI, non-White participants in the standardization sample of the PAI-A scored higher on the Paranoia Scale than White participants. Additional studies on the utility, validity, and reliability of the PAI-A with diverse samples still need to be conducted.

Summary

The PAI was developed to be a psychometrically sound and unbiased measure of personality. The test items were evaluated for potential racial bias by a review panel before a final set of items was agreed upon. African Americans were well represented in both the community and clinical standardization samples. Studies show that African Americans and Caucasians generally obtain similar scores on this measure, but African Americans tend to get higher scores on the Paranoia Scale. In order to make relative comparisons, separate Black norms based on the scores of the Black community sample are available. However, T scores derived from the full normative sample are recommended. The adolescent version of this measure, the PAI-A, still needs to be studied with diverse groups. Overall, the research on the PAI supports its use with African American samples; however, clinicians should discuss high scores on the Paranoia scale with their clients to ensure that they reflect clinical symptoms and not factors associated with racism.

The NEO Personality Inventories

The Revised NEO Personality Inventory (NEO PI-R)

The NEO PI-R (Costa and McCrae 1992) was developed to be a concise measure of the five major domains of personality (neuroticism, extraversion, openness, agreeableness, and conscientiousness), and in contrast to the other measures covered in this chapter, is not designed purely as a measure of pathology. The Five-Factor Model of personality is based on years of research and defines personality as endogenous basic tendencies within a specific cultural context (McCrae 2001). Longitudinal studies have demonstrated that personality traits tend to be fairly stable across the lifespan despite the influence of major life events (McCrae et al. 2000). The NEO PI-R is comprised of five scales for each of the five personality domains and six subscales for facets within each domain (Costa and McCrae 1992). Scores on the NEO PI-R are expressed in T scores, with most individuals scoring near the average for a particular scale. The scores are interpreted on a dimension, as either very low (T below 35), low ($T=35-45$), average ($T=45-55$), high ($T=55-65$), and very high (T above 65). It contains 240 items answered on a 5-point scale.

The normative sample of the NEO PI-R was comprised of 500 men and 500 women from nonclinical populations (Costa and McCrae 1992). The percentage of African individuals in the normative sample was equivalent to the US Census projections for 1995 (11.2 % African American males and 12.0 % African American females). Separate norms for African Americans are not available.

NEO PI-R Studies

The NEO PI-R scores of men and women from 26 different cultures were examined and compared to the American normative sample in a large study with 23,031 adults (McCrae 2001). The patterns of age- and gender-related differences on the NEO PI-R for the non-American samples were similar to those found in American samples. Subsamples within the same culture exhibited similar levels of personality traits. An intercultural factor analysis revealed a close approximation to the Five-Factor Model. Furthermore, the factor scores were meaningfully related to other conceptually relevant variables, indicative of convergent validity across cultures.

The NEO PI-R was administered to 141 African American and 92 Caucasian volunteer men and women and it was determined that the participant's education, household income, and father's and mother's education each had significant main effects on personality (Jonassaint et al. 2011). A pattern emerged such that high participant's socioeconomic status (SES)/high mother's education was associated with high extraversion and openness and low participant's SES/low father's education was associated with high neuroticism and low conscientiousness. These findings were consistent for both African American and Caucasian participants, suggesting that there are no racial differences in the impact of socioeconomic status on personality based on the NEO-PI-R.

The factor structure of the NEO PI-R was examined in a sample of 408 South African college students, 92 of which were Black (Heuchert et al. 2000). The authors found that the structure of the Five-Factor Model of the entire sample, and those of the Caucasian and Black subgroups, were consistent with that obtained with the American normative sample. However, there were some differences in personality scores between the racial groups. Caucasians scored significantly higher than Blacks on openness, extraversion, and agreeableness. The authors hypothesize that these differences are likely due to social, economic, and cultural differences between the racial groups.

The NEO Personality Inventory-3 (NEO PI-3)

The NEO PI-R was designed for use with adults and younger populations have demonstrated difficulty understanding some of the test items (Costa and McCrae 1992; McCrae et al. 2005a). Indeed, some adults have had difficulty understanding some of the wording, such as "fastidious and lackadaisical." To address this problem, the test developers created a revised version of the NEO PI-R, the NEO PI-3 (McCrae et al. 2005a). The NEO PI-3 has a fifth-grade reading level. In this newer version, 37 items that were determined to be difficult to understand by study participants were reworded to be more readable. The NEO PI-3 appears to have the same factor structure as the NEO PI-R and slightly better internal consistency. It has been shown to be a valid and reliability measure of personality in children as young as 12 years old (Costa et al. 2008). Separate norms are available for adolescents (12–20 years) and adults (21 years or older); however, the authors recommend using the combined-aged norms unless comparisons relative to peers are desired (McCrae et al. 2005b).

The NEO PI-3 was evaluated using data from observer ratings of adolescents from 24 different cultures from African, Asian, European, and North American and South American (De Fruyt et al. 2009). The investigators found that the NEO PI-3 maintained the structure of the Five-Factor Model and was psychometrically equivalent or slightly better than the NEO PI-R. More studies on the reliability and validity of the NEO PI-3 are still needed.

Summary

The NEO inventories are measures of normal personality domains and facets, rather than psychopathology. It is predominantly used with nonclinical populations. Cross-cultural studies across the globe have demonstrated that individuals from vastly different cultures obtain similar score patterns and factor structures. There have been no studies that specifically evaluated the validity of this measure with African Americans or compared the scores of African Americans with Caucasian Americans. The NEO PI-3 is recommended for adults with lower reading levels, adolescents, and children as young as 12 years of age. The evidence suggests that the NEO inventories are a valid measure of normal personality traits across cultures and. Clinicians should be aware that the results of the NEO represent an individual's endogenous traits within a cultural context, so racial factors will influence how the individual responds.

Summary and Recommendations

A lack of clinically meaningful differences between groups suggests that the personality measure is useful for assessing psychopathological constructs that exist across cultures. Most of the empirical studies on self-report personality instruments have been conducted with clinical samples, but when nonclinical samples were utilized it was more likely that significant differences were found between racial and ethnic groups, suggesting that cultural or adjustment-related factors may influence personality profiles (Kwan and Maestas 2008). More research in this area is needed before conclusions can be made. Research to identify expected group differences could help make personality assessments more useful and help control any potential test bias.

One of the first steps in choosing a self-report personality measure is assuring that the test is appropriate for the individual being tested (e.g., clinical vs. nonclinical, reading level). The ideal measure will have a normative sample that is well-represented with individuals who are demographically similar to the client in need of testing. The available research generally supports the use of the MMPI-2, MCMI-III, PAI, and NEO inventories with African American populations. However, several group differences have been demonstrated between African Americans and Caucasians when using these measures. The following are recommendations based on the available research:

The MMPI-2 has good psychometric support with African American populations, with most differences between groups no longer significant after considering the effects of sociodemographic variables. However, clinicians should examine profiles carefully and use other sources of information to corroborate findings because there is evidence that the MMPI-2 may underpredict psychopathology in African Americans. Some scales are better supported for use with African Americans than others. For example, the research does not support the use of the MAC-R with African Americans.

Research examining the use of the MCMI-III with diverse populations is conspicuously absent. The extant research does not strongly support its use with African Americans. Therefore, it is recommended that clinicians utilize one of the other personality measures until more research with African Americans is completed.

The PAI was clearly designed to address cultural bias from its inception and demonstrates few differences between African American and Caucasian profiles. Therefore, the PAI appears to be one of the best measures of personality for African American populations.

The NEO inventories are well studied with many cultures around the world and appear to be valid measures of non-pathological personality traits across different cultures. These measures were developed based on a theory that anticipants the influence of contextual factors on personality; therefore, scores will reflect cultural influences.

Test Interpretation

Some personality profile differences between African Americans and Caucasians may be due to personality distinctions associated with differential experiences between groups, such as racism, oppression, and racial socialization. These experiences could cause an individual to have a particular personality profile, such as high levels of paranoia or cynicism. The relationship between cultural variables and personality assessment test scores is still not well understood, which could result in inaccurate test interpretation. The more similar a client is to participants in the normative sample, the easier it will be to interpret scores. This will allow for scores to be interpreted in a similar way as the same score obtained by the normative sample. Scale elevations on the personality assessments of African Americans should be interpreted cautiously and with cultural awareness. Based on evidence that acculturation may influence how African Americans respond to items on personality assessments, acculturation measures and an examination of responses on individual items could help clinicians interpret personality profiles more accurately. An examination of individual items that lead to scale elevation could help clinicians understand if a scale elevation is due to actual personality characteristics and symptomatology or cultural factors that do not warrant clinical attention. In all cases, attempts should be made to gather

Table 4.1 At-a-glance summary table of the measures discussed in the chapter

Assessment name	Disorder assessed	Recommendations and relevant research findings
Minnesota Multiphasic Personality Inventory-2 (MMPI-2)	Personality/Psychopathology	Well researched in various population samples African Americans tend to score higher than Caucasians on various scales, especially 6, 8, 9, FRS, CYN, ASP, BIZ, but less so after controlling for demographics and diagnoses African Americans may score higher on validity scales, but this could be due to educational differences MAC-R may not be a valid measure of substance abuse for African Americans Overall, the MMPI-2 may underpredict pathology for African Americans
Millon Clinical Multiaxial Inventory-III (MCMI-III)	Personality/Psychopathology	Only for use with clinical populations Not well studied since the 1990's, with only one study since the publication of MCMI-III African Americans not well represented in test development and standardization samples African Americans tend to elevate the several scales, especially the Narcissistic, Aggressive, Paranoia, Drug, and Psychotic Delusion scales, but have similar factor structures as Caucasians The validity of the Drug Abuse scale with African Americans is questionable
Personality Assessment Inventory (PAI)	Personality/Psychopathology	The test items were evaluated for racial bias by a review panel African Americans were well represented in both community and clinical standardization samples African Americans tend to score higher than Caucasians on the Paranoia scale Separate Black norms are available for the purpose of making relative comparisons Research supports the validity and reliability of the PAI with African Americans
NEO Inventories (NEO PI-R, NEO PI-3)	Personality	The NEO is used predominantly with nonclinical populations It has been evaluated globally and shows similar score patterns and factor structures across cultures The NEO PI-3 is recommended for those with lower reading levels

corroborating information through additional tests and collateral sources. If additional, well-validated sources contradict scores on a measure of personality, it is possible that a different measure should be considered.

An examination of the extant research on personality assessment reveals that there has been a lack of attention to assessment of diverse groups, making it difficult to make empirically informed recommendations regarding the use of personality assessments with African Americans. A review of 60 years of personality assessment research demonstrated that only 6% of participants were identified as ethnic or racial minorities (Holaday and Boucher 1999). There continues to be a lack of attention to racial and cultural issues relevant to assessing personality with self-report measures. Future research will hopefully help psychologists make more accurate decisions when assessing personality in African American populations (Table 4.1).

References

Al-Halabí, S., Herrero, R., Saiz, P. A., Garcia-Portilla, M. P., Corcoran, P., Bascaran, M. T., et al. (2010). Sociodemographic factors associated with personality traits assessed through the TCI. *Personality and Individual Differences, 48,* 809–814. doi:10.1016/j.paid.2010.02.001.

Alterman, A. I., Zaballero, A. R., Lin, M. M., Siddiqui, N., Brown, L. S., Rutherford, M. J., & McDermott, P. A. (1995). Personality Assessment Inventory (PAI) scores of lower-socioeconomic African American and Latino methadone maintenance patients. *Assessment, 2,* 91–100. doi:10.1177/107319119500200109.

Alterman, A. I., Snider, E. C., Cacciola, J. S., Brown, L. S., Zaballero, A., & Siddiqui, N. (1996). Evidence for response set effects in structured research interviews. *The Journal of Nervous and Mental Disease, 184,* 403–410. doi:10.1097/00005053-199607000-00002.

Aranda, M. P., Chae, D. H., Lincoln, K. D., Taylor, R. J., Woodward, A. T., & Chatters, L. M. (2012). Demographic correlates of DSM-IV major depressive disorder among older African Americans, Black Caribbeans, and non-Hispanic Whites: Results from the National Survey of American Life. *International Journal of Geriatric Psychiatry, 27,* 940–947. doi:10.1002/gps.2805.

Arbisi, P. A., Ben-Porath, Y. S., & McNulty, J. (2002). A comparison of MMPI-2 validity in African American and Caucasian psychiatric inpatients. *Psychological Assessment, 14,* 3–15. doi:10.1037//1040-3590.14.1.3.

Barnes, A. (2008). Race and hospital diagnoses of schizophrenia and mood disorders. *Social Work, 53,* 77–83. doi:10.1093/sw/53.1.77.

Ben-Porath, Y. S., Shondrick, D. D., & Stafford, K. P. (1995). MMPI-2 and race in a forensic diagnostic sample. *Criminal Justice and Behavior, 22,* 19–32. doi:10.1177/0093854895022001002.

Boyd-Franklin, N., & Karger, M. (2012). Intersections of race, class, and poverty: Challenges and resilience in African American familes. In F. Walsh (Ed.), *Normal family processes: Growing diversity and complexity* (4th ed., pp. 273–296). New York: Guilford.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellengen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellengen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *Minnesota Mulitiphasic Personality Inventory-2: Manual for administration, scoring, and interpretation, revised edition*. Minneapolis: University of Minnesota Press.

Calsyn, D. A., Saxon, A. J., & Daisy, F. (1991). Validity of the MCMI Drug Abuse Scale varies as a function of the drug choice, race, and Axis II subtypes. *American Journal of Drug and Alcohol Abuse, 17,* 153–159. doi:10.3109/00952999108992818.

Castro, Y., Gordon, K. H., Brown, J. S., Anestis, J. C., & Joiner, T. E. (2008). Examination of racial differences on the MMPI-2 clinical and restructured clinical scales in an outpatient sample. *Assessment, 15,* 277–286. doi:10.1177/1073191107312735.

Choca, J. P., Shanley, L. A., Peterson, C. A., & Denburg, E. Van. (1990). Racial bias and the MCMI. *Journal of Personality Assessment, 54,* 479–490. doi:10.1080/00223891.1990.9674013.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R: Professional manual*. Lutz: Psychological Assessment Resources.

Costa, P. T., McCrae, R. R., & Martin, T. A. (2008). Incipient adult personality: The NEO-PI-3 in middle-school-aged children. *British Journal of Developmental Psychology, 26,* 71–89. doi:10.1348/026151007X196273.

Craig, R. J., & Bivens, A. (1998). Factor structure of the MCMI-III. *Journal of Personality Assessment, 70,* 190–196. doi:10.1207/s15327752jpa7001_13.

Cross, W. E. (1978). The Thomas and Cross models of psychological nigrescence: A review. *Journal of Black Psychology, 5,* 13–31. doi:10.1177/009579847800500102.

Dana, R. H., & Whatley, P. R. (1991). When does a difference make a difference? MMPI scores and African-Americans. *Journal of Clinical Psychology, 47,* 400–406. doi:10.1002/1097-4679(199105)47:3<400::AID-JCLP2270470312>3.0.CO;2-4.

Davis, W. E. (1975). Race and the differential “power” of the MMPI. *Journal of Personality Assessment, 39,* 138–140. doi:10.1207/s15327752jpa3902_8.

Davis, W. E., Greenblatt, R. L., & Pochyly, J. M. (1990). Test of MCMI Black norms for five scales.

Journal of Clinical Psychology, 46, 175–178. doi:10.1002/1097-4679(199003)46:2<175::AID-JCLP2270460208>3.0.CO;2–Q.

De Fruyt, F., De Bolle, M., McCrae, R. R., Terracciano, A., Costa, P. T., & Collaborators of the Adolescent Personality Profiles of Cultures Project. (2009). Assessing the universal structure of personality in early adolescence: The NEO-PI-R and NEO-PI-3 in 24 cultures. *Assessment, 16,* 301–311. doi:10.1177/1073191109333760.

Dean, A. C., Boone, K. B., Kim, M. S., Curiel, A. R., Martin, D. J., Victor, T. L., et al. (2008). Examination of the impact of ethnicity on the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) Fake Bad Scale. *The Clinical Neuropsychologist, 22,* 1054–1060. doi:10.1080/13854040701750891.

Denavas-walt, C., Proctor, B. D., Smith, J. C., & U.S. Census Bureau. (2012). Current population reports, P60-243. *Income, poverty, and health insurance coverage in the United States: 2011*. http://www.census.gov/prod/2012pubs/p60-243.pdf. Accessed 1 June 2013.

Frueh, B. C., Smith, D. W., & Libet, J. M. (1996). Racial differences on psychological measures in combat veterans seeking treatment for PTSD. *Journal of Personality Assessment, 66,* 41–53. doi:10.1207/s15327752jpa6601_3.

Frueh, B. C., Gold, P. B., de Arellano, M. A., & Brady, K. L. (1997). A racial comparison of combat veterans evaluated for PTSD. *Journal of Personality Assessment, 68,* 692–702. doi:10.1207/s15327752jpa6803_14.

Ghafoori, B., & Hierholzer, R. W. (2010). Personality patterns among Black, White, and Hispanic combat veterans. *Psychological Trauma: Theory, Research, Practice, and Policy, 2,* 12–18. doi:10.1037/a0019019.

Gibbs, T. A., Okuda, M., Oquendo, M. A., Lawson, W. B., Wang, S., Thomas, Y. F., & Blanco, C. (2013). Mental health of African Americans and Caribbean Blacks in the United States: Results from the National Epidemiological Survey on alcohol and related conditions. *American Journal of Public Health, 103,* 330–338. doi:10.2105/AJPH.2012.300891.

Goldman, V. J., Cooke, A., & Dahlstrom, W. G. (1995). Black-White differences among college students: A comparison of MMPI and MMPI-2 norms. *Assessment, 2,* 293–299. doi:10.1177/1073191195002003009.

Gómez, F. C., Johnson, R., Davis, Q., & Velásquez, R. J. (2000). MMPI-A performance of African American and Mexican American adolescent first-time offenders. *Psychological Reports, 87,* 309–314. doi:10.2466/PR0.87.5.309-314.

Graham, J. R. (2006). *MMPI-2: Assessing personality and psychopathology*. New York: Oxford University Press.

Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan, W. J., & Pickering, R. P. (2004). Prevalence, correlates, and disability of personality disorders in the United States: Results from the national epidemiologic survey on alcohol and related conditions. *Journal of Clinical Psychiatry, 65,* 948–958.

Greenblatt, R. L., & Davis, W. E. (1992). Accuracy of MCMI classification of angry and psychotic Black and White patients. *Journal of Clinical Psychology, 48,* 59–63. doi:10.1002/1097-4679(199201)48:1<59::AID-JCLP2270480108>3.0.CO;2-Z.

Greene, R. L. (1987). Ethnicity and MMPI performance: A review. *Journal of Consulting and Clinical Psychology, 55,* 497–512. doi:10.1037/0022-006X.55.4.497.

Gripshover, D. L., & Dacey, C. M. (1994). Discriminative validity of the MacAndrew Scale in settings with a high base rate of substance abuse. *Journal of Studies on Alcohol, 55,* 303–308.

Hall, G. C. N., & Phung, A. H. (2001). Minnesota Multiphasic Personality Inventory and Millon Clinical Multiaxial Inventory. In L. A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 307–330). San Francisco: Jossey-Bass.

Hall, G. C. N., Bansal, A., & Lopez, I. R. (1999). Ethnicity and psychopathology: A meta-analytic review of 31 years of comparative MMPI/MMPI-2 research. *Psychological Assessment, 11,* 186–197. doi:10.1037//1040-3590.11.2.186.

Hamberger, L. K., & Hastings, J. E. (1992). Racial differences on the MCMI in an outpatient clinical sample. *Journal of Personality Assessment, 58,* 90–95. doi:10.1037/a0019019.

Harrell, S. P. (2000). A multidimensional conceptualization of racism-related stress: Implications for the well-being of people of color. *American Journal of Orthopsychiatry, 70,* 42–57. doi:10.1037/h0087722.

Harrison, R. H., & Kass, E. H. (1967). Differences between Negro and White pregnant women on the MMPI. *Journal of Consulting Psychology, 31,* 454–463. doi:10.1037/h0025012.

Heuchert, J. W. P., Parker, W. D., Stumpf, H., & Myburgh, C. P. H. (2000). The Five-Factor Model of personality in South African college students. *American Behavioral Scientist, 44,* 112–125. doi:10.1177/00027640021956125.

Holaday, M., & Boucher, M. (1999). Journal of personality assessment: 60 years. *Journal of Personality Assessment, 72,* 111–124. doi:10.1207/s15327752jpa7201_7.

Hutton, H. E., Miner, M. H., Blades, J. R., & Langfeldt, V. C. (1992). Ethnic differences on the MMPI Overcontrolled-Hostility scale. *Journal of Personality Assessment, 58,* 260–268. doi:10.1207/s15327752jpa5802_5.

Jonassaint, C. R., Siegler, I. C., Barefoot, J. C., Edwards, C. L., & Williams, R. B. (2011). Low life course socioeconomic status (SES) is associated with negative NEO PI-R personality patterns. *International Journal of Behavioral Medicine, 18,* 13–21. doi:10.1007/s12529-009-9069-x.

Klonoff, E. A., Landrine, H., & Ullman, J. B. (1999). Racial discrimination and psychiatric symptoms among Blacks. *Cultural Diversity & Ethnic Minority Psychology, 5,* 329–339. doi:10.1037/1099-9809.5.4.329.

Kwan, K.-L. K., & Maestas, M. L. (2008). MMPI-2 and MCMI-III performances of non-White people in the United States: What we (don't) know and where we go from here. In Lisa A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical,* *psychological, and educational applications* (pp. 425–446). San Francisco: Jossey-Bass.

Long, K. A., Graham, J. R., & Timbrook, R. E. (1994). Socioeconomic status and MMPI-2 interpretation. *Measurement and Evaluation in Counseling and Development, 27,* 158–177.

Lynn, R. (2002). Racial and ethnic differences in psychopathic personality. *Personality and Individual Differences, 32,* 273–316. doi:10.1016/S0191-8869(01)00029-0.

McCrae, R. R. (2001). Trait psychology and culture: Exploring intercultural comparisons. *Journal of Personality, 69,* 819–846. doi:10.1111/1467-6494.696166.

McCrae, R. R., & Terracciano, A. (2006). National character and personality. *Current Directions in Psychological Science, 15,* 156–161. doi:10.1111/j.1467-8721.2006.00427.x.

McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hrebícková, M., Avia, M. D., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78,* 173–186. doi:10.1037//0022-3514.78.1.173.

McCrae, R. R., Costa, P. T., & Martin, T. A. (2005a). The NEO-PI-3: A more readable revised NEO Personality Inventory. *Journal of Personality Assessment, 84,* 261–270. doi:10.1207/s15327752jpa8403_05.

McCrae, R. R., Martin, T. A., & Costa, P. T. (2005b). Age trends and age norms for the NEO Personality Inventory-3 in adolescents and adults. *Assessment, 12,* 363–373. doi:10.1177/1073191105279724.

McCrae, R. R., Terracciano, A., & 79 Members of the Personality Profiles of Cultures Project. (2005c). Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology, 89,* 407–25. doi:10.1037/0022-3514.89.3.407.

McNulty, J. L., Graham, J. R., Ben-Porath, Y. S., & Stein, L. A. R. (1997). Comparative validity of MMPI-2 scores of African American and Caucasian mental health center clients. *Psychological Assessment, 9,* 464–470. doi:10.1037/1040-3590.9.4.464.

McNulty, J. L., Forbey, J. D., Graham, J. R., Ben-Porath, Y. S., Black, M. S., Anderson, S. V., & Burlew, A. K. (2003). MMPI-2 validity scale characteristics in a correctional sample. *Assessment, 10,* 288–298. doi:10.1177/1073191103255623.

Micucci, J. A. (2002). Accuracy of MMPI-A scales ACK, MAC-R, and PRO in detecting comorbid substance abuse among psychiatric inpatients. *Assessment, 9,* 111–122. doi:10.1177/10791102009002001.

Millon, T. (1984). *Millon Clinical Multiaxial Inventory manual supplement*. Minneapolis: National Computer System.

Millon, T. (1993). *The Millon Adolescent Clinical Inventory*. Minneapolis: National Computer System.

Millon, T. (1994). *MCMI-III manual*. Minneapolis: National Computer System.

Millon, T., Millon, C., Davis, R., & Grossman, S. (2006). *MCMI-III manual* (3rd ed.). Minneapolis: Pearson Assessments.

Millon, T., Millon, C., Davis, R., & Grossman, S. (2009). *MCMI-III manual* (4th ed.). Minneapolis: Pearson Assessments.

Monnier, J., Elhai, J. D., Frueh, B. C., Sauvageot, J. A., & Magruder, K. M. (2002). Replication and expansion of findings related to racial differences in veterans with combat-related PTSD. *Depression and Anxiety, 16,* 64–70. doi:10.1002/da.10060.

Monnot, M. J., Quirk, S. W., Hoerger, M., & Brewer, L. (2009). Racial bias in personality assessment: Using the MMPI-2 to predict psychiatric diagnoses of African American and Caucasian chemical dependency inpatients. *Psychological Assessment, 21,* 137–151. doi:10.1037/a0015316.

Morey, L. C. (1991). *The Personality Assessment Inventory: Professional manual*. Lutz: Psychological Assessment Resources.

Morey, L. C. (2007a). *The Personality Assessment Inventory: Professional manual* (2nd ed.). Odessa: Psychological Assessment Resources.

Morey, L. C. (2007b). *Personality Assessment Inventory-Adolescent*. Lutz: Psychological Assessment Resources.

Munley, P. H., Vacha-Haase, T., Busby, R. M., & Paul, B. D. (1998). The MCMI-II and race. *Journal of Personality Assessment, 70,* 183–189. doi:10.1207/s15327752jpa7001_12.

Munley, P. H., Morris, J. R., Murray, D. A., & Baines, T. C. (2001). A comparison of African American and White American veteran MMPI-2 profiles. *Assessment, 8,* 1–10. doi:10.1177/107319110100800101.

Parker, J. D., Daleiden, E. L., & Simpson, C. A. (1999). Personality Assessment Inventory substance-use scales: Convergent and discriminant relations with the Addiction Severity Index in a residential chemical dependence treatment setting. *Psychological Assessment, 11,* 507–513. doi:10.1037/1040-3590.11.4.507.

Regier, D. A., Farmer, M. E., Rae, D. S., Myers, J. K., Kramer, M., Robins, L. N., et al. (1993). One-month prevalence of mental disorders in the United States and sociodemographic characteristics: The Epidemiologic Catchment Area study. *Acta Psychiatrica Scandinavica, 88,* 35–47. doi:10.1111/j.1600-0447.1993.tb03411.x.

Schinka, J. A., LaLone, L., & Greene, R. L. (1998). Effects of psychpathology and demographic characteristics on MMPI-2 scale scores. *Journal of Personality Assessment, 70,* 197–211. doi:10.1207/s15327752jpa7002_1.

Shea, S. J., Mckee, G. R., Shea, M. E. C., & Culley, D. C. (1996). MMPI-2 profiles of male pre-trial defendants. *Behavioral Sciences and the Law, 14,* 331–338. doi:10.1002/(SICI)1099-0798(199622)14:3.

Skeem, J. L., Edens, J. F., Sanford, G. M., & Colwell, L. H. (2003). Psychopathic personality and racial/ethnic differences reconsidered: A reply to Lynn (2002). *Personality and Individual Differences, 35,* 1439–1462. doi:10.1016/S0191-8869(02)00361-6.

Stefurak, T., Calhoun, G. B., & Glaser, B. A. (2004). Personality typologies of male juvenile offenders using a cluster analysis of the Millon Adolescent Clinical Inventory introduction. *International Journal of Offender Therapy and Comparative Criminology, 48,* 96–110. doi:10.1177/0306624X03258478.

Stukenberg, K., Brady, C., & Klinetob, N. (2000). Psychiatric inpatients and the MMPI-2: Providing benchmarks. *Journal of Clinical Psychology, 56,* 747–756. doi:10.1002/(SICI)1097-4679(200006)56:6<747::AID-JCLP5>3.0.CO;2-O.

Timbrook, R. E., & Graham, J. R. (1994). Ethnic differences on the MMPI-2? *Psychological Assessment, 6,* 212–217. doi:10.1037//1040-3590.6.3.212.

U.S. Census Bureau. (2011a). Law enforcement, courts, and prisons. Statistical abstract of the United States. http://www.census.gov/prod/2011pubs/12statab/law.pdf. Accessed 1 June 2013.

U.S. Census Bureau. (2011b). The Black alone or in combination population in the United States. Annual Social and Economic (ASEC) Supplement to the Current Population Survey (CPS). http://www.census.gov/population/race/data/ppl-bc11.html. Accessed 1 June 2013.

U.S. Census Bureau. (2012a). Labor force, employment, and earnings. Statistical abstract of the United States. http://www.census.gov/compendia/statab/2012/tables/12s0588.pdf. Accessed 1 June 2013.

U.S. Census Bureau. (2012b). Educational attainment in the United States: 2011—detailed tables. Educational attainment. http://www.census.gov/hhes/socdemo/education/data/cps/2011/tables.html. Accessed 1 June 2013.

U.S. Census Bureau. (2013). USA QuickFacts. http://quickfacts.census.gov/qfd/states/00000.html. Retrieved 21 April 2013. Accessed 1 June 2013.

Walters, G. D., Greene, R. L., Jeffrey, T. B., Kruzich, D. J., & Haskin, J. J. (1983). Racial variations on the MacAndrew Alcoholism Scale of the MMPI. *Journal of Consulting and Clinical Psychology, 51,* 947–948. doi:10.1037/0022-006X.51.6.947.

Walters, G. D., Greene, R. L., & Jeffrey, T. B. (1984). Discriminating between alcoholic and nonalcoholic Blacks and Whites on the MMPI. *Journal of Personality Assessment, 48,* 486–488. doi:10.1207/s15327752jpa4805_6.

Whatley, P. R., Allen, J., & Dana, R. H. (2003). Racial identity and the MMPI in African American male college students. *Cultural Diversity & Ethnic Minority Psychology, 9,* 345–353. doi:10.1037/1099-9809.9.4.345.

Youman, K., Drapalski, A., Stuewig, J., Bagley, K., & Tangney, J. (2010). Race differences in psychopathology and disparities in treatment seeking: Community and jail-based treatment seeking patterns. *Psychological Services, 7,* 11–26. doi:10.1037/a0017864.

Zuckerman, M. (2003). Are there racial and ethnic differences in psychopathic personality? A critique of Lynn's (2002) racial and ethnic differences in psychopathic personality. *Personality and Individual Differences, 35,* 1463–1469. doi:10.1016/S0191-8869(02)00362-8.

IQ Testing and the African American Client

5

Nicholas S. Thaler, April D. Thames, Xavier E. Cagigas and Marc A. Norman

Background and History

Intelligence assessment has a rich history in the psychologists' role of professional duties that can be traced back to World War I. IQ scores have been shown to be a reasonably good predictor of grades at school, performance at work, and many other aspects of success in life (Gottfredson 2004; Herrnstein and Murray 1994). However, history has also demonstrated that when IQ group score differences are either ignored or poorly understood, there is danger for misinterpretation. Racial differences emerged as an explanation for intellectual performance differences between groups around the late nineteenth and early twentieth centuries, which in turn sparked large-scale debates regarding genes, race, and intelligence.

The manual of the Stanford–Binet Intelligence Test provides an illustrative example of how IQ scores can be misrepresented to substantiate racism. In this manual Lewis Terman, the psychologist and developer of the Stanford–Binet, noted performance discrepancies between white and non white American racial groups and stated that the "enormously significant racial differences in general intelligence could not be remedied by education" (Terman 1916). In 1930, William Shockley made a controversial argument regarding genetics and reproduction among those who were deemed to be of low intellect. Shockley argued that a dysgenic effect (i.e., promoting reproduction among those with less-intellect at the expense of those with high intellect) would ultimately lead to a drop in overall average intelligence that would negatively impact civilization. Although Shockley was concerned about dysgenic effects among both blacks and whites, he viewed this being more problematic among blacks.

Further controversy surrounding race and measured intelligence followed the 1994 publication of *The Bell Curve* by Richard Herrnstein and Charles Murray (Hernstein and Murray 1994). This work attempted to explain variations in intelligence in American society using statistical analyses, raise warnings regarding the consequences of this intelligence gap, and propose a national social policy with the goal of mitigating the worst of the consequences attributed to this intelligence gap. Many of the assertions and conclusions made by the authors were indeed controversial, including the notion that intelligence was the cause for antisocial behavior, and that genetics were responsible for the observed performance differences between African Americans when compared to Whites and Asians.

Conclusions such as those made in *The Bell Curve* have long since been debunked. For example, a review article published by the Journal

---

N. S. Thaler (✉) · A. D. Thames · X. E. Cagigas
Department of Psychiatry and Biobehavioral Sciences,
Semel Institute for Neuroscience and Human Behavior,
University of California, Los Angeles, 760 Westwood
Plaza, Los Angeles, CA 90024, USA
e-mail: nick.thaler@gmail.com

M. A. Norman
Department of Psychiatry, University of California,
San Diego 200 West Arbor Drive, San Diego,
CA 92103-8218, USA

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_5, © Springer Science+Business Media New York 2015

of the American Psychologist (Nisbett et al. 2012) provides evidence to support that IQ heritability differs as a function of social class and that there is virtually no evidence to support genetic polymorphisms associated with intelligence among individuals who score within the normal IQ range. Importantly, the review draws attention to the reduction in the IQ gap between African Americans and Whites and how environmental factors (such as stress) influence cognitive functioning. In this chapter, we further discuss environmental and psychosocial contributors to test performance that are particularly relevant to African Americans.

In practice, the onus is placed on the psychologist to understand the utility and limitations of IQ assessment with regard to how environmental, psychological, and sociological factors can influence scores. One of the critiques of IQ assessment is that it arises from a predominantly European–Western conceptualization of intellectual ability and does not consider emic perspectives of intelligence that are meaningful to specific cultural groups (Thomsen et al. 2008). IQ scores are influenced by social background, class, and academic achievement and are further impacted by test bias (Reynolds 2000) that often favor Westernized culture. Suboptimal performance can be influenced by a number of sources including psychometric limitations of test measures, issues pertaining to the assessment context, performance anxiety and test effort and motivation, and differential access to resources (e.g., quality education, test familiarity). All of these are threats to the overall construct of "intelligence."

Sources of Bias and Construct Validity Threats

We would be remiss not to provide a brief overview of the various sources of test bias and threats to construct validity that can occur in the context of testing. While these principles are more general to the assessment of individuals from minority and disadvantaged groups, they are particularly relevant to the African American client.

Measurement Bias

When certain test characteristics are not related to the construct being measured or the manner in which the test is used, in the probability of different meanings for scores earned by members of different identifiable subgroups is almost inevitable. Differential item functioning (DIF) is said to occur when test takers differ in their probabilities of answering a test item correctly as a function of group membership. DIF can be evaluated in a variety of ways. Groups must be matched on the measured characteristic, otherwise DIF may fail to be detected. Differential test functioning (DTF) refers to differences in the functioning of tests (or sets of items) for different specially defined groups. When DTF occurs, individuals from different groups who have the same standing on the characteristic assessed by the test do not have the same expected test score. The term predictive bias may be used when evidence is found that differences exist in the patterns of associations between test scores and other variables for different groups, bringing with it concerns about bias in the inferences drawn from the use of test scores.

Uniform DIF is present when one group is more likely to provide a correct response than another group across all levels of ability. A clear example of uniform DIF would the probability of an English speaker correctly defining a vocabulary word compared to a non-English speaker; in other words, regardless of intellectual ability, the non-English speaker is unable to accurately define an English word. Nonuniform DIF is present when the probability of a correct response varies across ability. For example, research on the Boston Naming Test has demonstrated that certain words (e.g., dominoes, tripod, palette) have nonuniform DIF between older whites and African Americans (Pedraza et al. 2009). DIF and DTF can be investigated through modern psychometric techniques such as item response theory (IRT) and contingency tables. These methods are more effective in detecting biased items compared to classical ANOVA methods, though the latter remains the dominant methodological approach (Reynolds 2000).

Another issue with tests is their psychometric equivalence across populations. The four-factor structure of the Wechsler Adult Intelligence Scales, third and fourth editions purports to represent four distinct IQ constructs. These constructs in turn are hypothesized to represent discrete strengths and weaknesses in cognition. However, the factor structure of tests depends on test performance and groups who do not perform in a similar fashion to the referent group may yield different factor structures (Allen et al. 2012; Delis et al. 2003). This begs the question "What are we actually measuring?" While psychometric equivalency studies have been published for some cultural groups (García and Abad 2003; Golay and Lecerf 2011), there remains a lack of such information in African American populations.

Access to Resources

With regard to intelligence testing, it is often assumed that all intended test takers have had equal opportunity to access the knowledge, skills, and abilities that reflect the measured construct. Accessibility is actually a test bias issue because obstacles to accessibility can result in different interpretations of test scores for individuals from different groups. For some test takers, factors related to individual characteristics such as age, socioeconomic status (SES), cultural background, disability, and/or English language proficiency may restrict accessibility and thus interfere with the measurement of the construct(s). Therefore, it is critical for the examiner to consider the extent to which individuals have had exposure to instruction or knowledge that affords them the opportunity to learn the content and skills targeted by the test. For example, an individual who has had little prior exposure to school may not have had the opportunity to learn concepts assumed to be common knowledge. It is important to note that exposure is not limited to the academic setting and can relate to cultural references, which are acquired through deep and chronic immersion in specific cultural practices. To illustrate this, Robert Williams demonstrated that African Americans outperformed whites on an intelligence test that was designed with the language and culture of African Americans as the initial point of reference (i.e., Black Intelligence Test of Cultural Homogeneity, or BITCH—100). This essentially provided a context in which familiarity with the African American culture resulted in better performance (Williams 1972). Failing to account for prior opportunity to learn the concepts tapped into by an IQ test can often lead to misdiagnosis, social stigma, inappropriate placement, and/or inappropriate assignment of services.

Test Content and Response Bias

When the actual test content differentially favors individuals from some subgroups over others, it creates construct-irrelevant variance. For example, a test that is intended to measure reading comprehension should not include words and expressions that are associated with particular occupations, disciplines, cultural backgrounds, SES, ethnic groups, or geographical locations. Appropriate test content should maximize the measurement of the latent construct and minimize confounding of this measurement with prior knowledge and experience that are likely to advantage, or disadvantage, test takers from particular subgroups.

Test response bias is thought to occur when items are solved in ways that were not originally intended. To the extent that such responses are more typical of some subgroups than others, biased score interpretations may result. For example, some cultural groups prefer to approach a task in a slow and cautious manner to ensure accuracy, which would compromise their score on timed tasks. There are specific cultural values that differ among whites and African Americans, which in turn often affect the cognitive approaches to assessment employed by psychologists. For instance, successful performance on several IQ measures relies heavily upon speed, and the African American culture generally values a cognitive style in which speed may be compromised in favor of accuracy (Byrd et al. 2004; Diehr et al.

1998).Similarly, different values associated with the nature and degree of verbal output can influence test taker responses. Some individuals may judge verbosity or rapid speech as rude, whereas others may regard those speech patterns as indications of high mental ability or friendliness. Obviously these contextual performance factors in assessment reflect broader cultural practices that, although frequently observed in African Americans, clearly represent motivational and dispositional factors toward testing that are more widespread in nature across other historically underrepresented cultural, linguistic, and ethnic groups. This highlights the need to move beyond simple racial and/or ethnic membership toward actually identifying the relevant cultural practices that foster, organize, and maintain these dispositional factors toward specific dimensions of cognitive assessment practices.

Still others may use a dialect that is not typically used in the dominant culture. For example, a common misconception about English is that it is a fixed language that is essentially homogeneous within the USA (Hilliard 1997). Such a misconception can bias performance on vocabulary tests, which often do not account for variations in English and penalize respondents who do not provide definitions in a “mainstream” manner. Subcultures within the USA, including those within the African American community demonstrate considerable variability in their use of the English language that are deeply contextual and rooted in specific cultural practices. This in turn may affect performance on IQ tests in which the content of a verbal response is judged and scored by the examiner from a more “mainstream” criterion base that is not inclusive of the culture-specific dimensions of language use.

Quality vs. Quantity of Education

Research has linked performance on IQ tests to a number of factors including access to resources, health literacy, social privilege, normative expectations of performance, all of which can result in performance discrepancies between African Americans and whites on IQ tests (Anglin and Kwate 2009). Turkheimer et al. (2003) conducted an analysis of SES by heritability interactions in the National Collaborative Perinatal Project. Structural equation modeling demonstrated large and statistically significant interactions between IQ and SES for Full Scale and Performance IQ (PIQ) but not for Verbal IQ (VIQ), although a trend for VIQ was in the same direction. For families at the lowest levels of SES, shared environment accounted for almost all of the variation in IQ, with genes accounting for practically none.

Discrepancies in performance scores, however, are not fully accounted for by equating demographic variables such as years of education or SES across groups (Manly et al. 2002). Matching groups on such metrics is an oversimplified attempt to control for group differences that neglects important psychosocial aspects. For example, while many studies match for years of education, other studies have reported that *quality* of education is a more important influence on IQ test performance (Manly et al. 2002, 2004). For example, discrepancies in quality of education likely reflect the fact that older cohorts were subjugated to racial segregation that confined African Americans to poorer educational environments. Despite the 1954 *Brown v. Board of Education* decision and several decades of desegregation efforts, younger African American generations are still more likely to be placed in school systems that have access to fewer educational resources and quality teachers than those in middle-class (and often white) neighborhoods. The Coleman report (Coleman 1966) was among the first of papers to demonstrate that quality of education can account for differences in outcomes between African Americans and whites and these findings have since been extensively replicated (Hedges et al. 1994; Manly et al. 2002; Manly et al. 2004; O’Neill 1990). In essence, quality of education as a construct is often deeply steeped in socioeconomic disparities driven by very real sociological factors such as the degree of segregation in neighborhoods, and institutionalized factors that drive social inequalities in school systems (e.g., funding sources being driven by property taxes within a zoned area). More directly stated, our assessment instruments are likely diluted by capturing variability that is perhaps more related to social inequality and other biasing factors reflective of the social context and are, therefore, not likely pure measures of the purported construct of interest that is trying to be measured.

Pediatric Assessment

As the research demonstrates, quality of education has a profound impact on the IQ and neuropsychological performance of African American adults. Quality of education is equally (if not more) important than years of education to account for when assessing African American children. Children who are not part of the majority culture are more likely to be referred for psychoeducational or special needs evaluations and are subsequently more likely to be diagnosed with intellectual and learning disabilities (Marlowe 2000). African American children specifically have faced discrimination through IQ testing, which disproportionately placed children in special education classes (Affeldt 2000). In California, this was addressed by the Larry P vs. Riles court case in 1979 that specifically prohibited using standardized IQ tests as the sole method of diagnosing African American children with mental retardation. Despite this case ruling, the debate remains on whether or not IQ batteries are valid methods of assessing ability in African American children. IQ batteries designed for children are subjected to the same sources of bias as adult measures (i.e., paucity of empirical research on the psychometric properties and psychosocial issues) that may explain performance discrepancies observed in African American children.

Children from impoverished backgrounds face the challenges associated with family poverty, access to health care and proper nutrition, and poorer community support. These children are at greater risk for being exposed to family distress due to limited finances, unsafe living conditions, absent family members, and poorer parental education and occupational status. Unfortunately, African American children, particularly urban youth, continue to represent a disproportionate number of individuals from such backgrounds. These psychosocial factors can impact the neurodevelopment of all children and as such, account for some of the IQ discrepancies observed between black and white children (Brazziel 2000). In addition, other sources of bias persist. As with adults, differences in familiarity with the dominant language impact children's performance on standardized tests. Parental involvement can also impact test performance; familial beliefs about illnesses and disability can impact a child's motivation and effort towards testing (Marlowe 2000). Another common criticism of IQ tests for children is a lack of norms for minority groups. African American children continue to be underrepresented in standardization samples and research on sources of content, measurement, and test bias remain lacking.

Demographically Adjusted Norms

Researchers have developed statistical corrections for score differences. Clinical Neuropsychology, a discipline that relies heavily upon assessment, has made considerable efforts to develop demographically corrected racial norms for several neuropsychological and IQ measures (e.g., Heaton et al. 2003). These norms were designed to equate scores among racial groups to minimize misclassifying individuals as cognitively impaired. More recently, computerized scoring programs for the WAIS-III and IV (i.e., Advanced Clinical Solutions: ACS) provide an in-depth method of psychometrically adjusting scores based on demographic differences that include race.

Because of observed racial score differences, African Americans are more likely to be misclassified as cognitively impaired when normative testing procedures are used (Manly et al. 1998). Therefore, demographic corrections provide a method of interpreting age-scaled test data that boosts specificity for diagnostic purposes (Heaton et al. 2001). While we acknowledge the clinical advantages of adjusted norms for diagnostic purposes, there are several limitations inherent to adjusted norms that warrant attention. (Manly 2005; Reitan and Wolfson 2005).

Principally, separate racial norms may promote individual biases and misinformation as to the *reasons* for having such norms—clinicians unfamiliar with the specifics of cultural neuropsychology may take on a reductionist perspective on the use of norms to adjust for a group that simply "performs poorer." Second, racial and ethnic categories are not tied to actual neurological or neurocognitive differences but rather serve as proxy for other variables of interest such as SES, quality and exposure to education in mainstream culture, and level of acculturation. In other words, a person's racial or ethnic classification provides little to no information about educational background, language proficiency, SES, acculturation, racial/ethnic identity, or the actual cultural practices people engage in which bring meaning and organization to their everyday lives (Cagigas and Manly 2014). Therefore, such classifications at best are a "short-hand" method of allowing for group corrections.

From a psychometric perspective, many corrected norms are derived from multiple regression that adjust age-scaled scores. While such methods are certainly acceptable ways to improve diagnostic specificity, in the end, the naïve consumer may believe that this ultimately solves the compounded problem of interpretability. In fact normative comparisons only provide a sense of how typical a particular score is relative to a reference group, but tell us very little about the underlying construct being assessed. In other words, such adjustments try to anchor the parameters that would otherwise lead to misclassification, misdiagnosis, and skewed interpretation, but do little to explain how a specific cognitive domain is expressed within the individual's culture, which raises the deeper question of construct validity; what do the tests actually measure cross-culturally and is this the same across groups? (Cagigas and Manly 2014). Further, demographically corrected norms are perhaps most appropriate in settings where the primary goal is to identify neurocognitive impairment. The use of demographic adjustments in the context of educational and vocational assessments may be more controversial, as the primary purpose of these types of assessments is to qualify an individual's ability to function in the school/work environment. Finally, adjusted norms and cutoffs might lose sensitivity to the presence of actual clinical impairment, particularly when misused with a nonrepresentative sample to draw inferences regarding an individual's level of functioning. For example, an African American male with advanced education who exhibits actual neurological impairment may be erroneously diagnosed as cognitive well preserved after using poorly-developed norms, as any actual weaknesses would be obscured by artificial inflation of test scores. However, recent evidence suggests that when properly developed, demographically corrected norms do not artificially inflate scores (Norman et al. 2011), but rather can improve diagnostic specificity when properly developed, applied, and interpreted.). As such, demographically-corrected norms can be a powerful and important tool in improving diagnostic specificity when in the hands of a properly trained clinician.

As mentioned previously, quality of education is thought to be a more precise method of estimating test performance above and beyond standard years of education. Considering that educational quality may be difficult to ascertain without a detailed knowledge of the patient's schooling, a proxy that is commonly used is reading proficiency. Reading proficiency tests require examinees to sound out regular and irregular words and individuals who were exposed to a better quality of mainstream education are more likely to be familiar with and pronounce the irregular words. Manly et al. (2002) demonstrated that controlling for literacy reduced score differences between older African Americans and whites on most neuropsychological tests. Dotson et al. (2008) demonstrated that years of education did not add to neuropsychology test score variance beyond literacy levels, which led to the development of literacy-based norms for lower SES African Americans. Acculturation, a construct that is typically studied in immigrant populations, is also demonstrably affiliated with test differences between African Americans and whites (Manly et al. 2004). Assessing acculturation has its own challenges. Broadly designed questionnaires may provide convenience but lack ecological validity for specific groups, while measures designed for specific groups do not generalize well and have a risk of being misused. For example, is an African American acculturation measure appropriate for use on a recent immigrant from the Ivory Coast? Within the USA itself there are different subcultures, and measures tailored to very narrow groups (e.g., Louisiana Creole) may have limited research and general applicability.

Stereotype Threat

One well-documented risk factor for poorer test performance is stereotype threat, or the fear of being judged based on a negative stereotype of a person's specific group. Research has consistently documented the effects of this phenomenon on test performance in African Americans and studies that control for stereotype threat report that score discrepancies between racial groups are drastically reduced or eliminated entirely. In the original work by Steele and Aronson (1995), African American and European American college students from Stanford University were randomly assigned to either an experimental "diagnostic" condition or a control condition. In the experimental/diagnostic condition, participants were told that they would be given a test that reflected intellectual capacity. The control group, however, was told that they would be participating in an "exercise." African Americans in the experimental condition, in which the stereotype of underperformance was activated, performed dramatically worse compared to African Americans in the control group. No performance differences were observed across ethnicities in the control group. The stereotype threat effect has been replicated with stereotypes related to age and gender (Hess et al. 2003; Spencer et al. 1999), and even golf performance (Stone et al. 1999). Although stereotype threat has been well documented in studies of IQ, math, and work performance (Steele and Aronson 1995; Ben-Zeev et al. 2005; Jamieson and Harkins 2007), the effect of stereotype threat on neuropsychological performance has received little attention until recently (Kit et al. 2014; Thames et al. 2013). In a nonclinical sample of African American and European adults, participants were randomly assigned to a stereotype threat or nonthreat condition (identical to the experimental manipulations outlined in Steele and Aronson 1995). Within each study condition, participants were assigned to either a same race or different race examiner and completed a measure of perceived discrimination and underwent neuropsychological testing. African Americans in the stereotype threat condition performed significantly worse on global neuropsychological functioning than African Americans in the nonthreat condition. African Americans who reported high levels of perceived discrimination performed significantly worse on memory tests when tested by an examiner of a different race, than African Americans who were tested by an examiner of the same race (Thames et al. 2013).

Proactively preparing participants who may experience possible stereotype threat has been demonstrated to improve outcomes in African American youth (Cohen et al. 2006). A detailed review on the mechanisms underlying stereotype threat extends beyond the scope of this chapter though briefly, values attributed to the testing domain of interest are an important moderator of who is vulnerable to stereotype threat (Steele 1997). In other words, individuals who have low domain identity towards testing and academic achievement may express this through poorer motivation and effort. Specific to African Americans, there has been recent research suggesting that racial identity is a moderator of stereotype effects within this group. For example, Davis et al. (2006) and McFarland et al. (2003) both reported that African Americans with a secure sense of racial identity outperformed those with a weaker racial identity. Finally, individual perception of stereotype threat may play a role in moderating test performance (Ployhart et al. 2003).

Related to the concept of stereotype threat, African Americans compared to European Americans and other racial/ethnic minorities demonstrate higher levels of test-taking anxiety. For example, studies have shown that African American students score higher on measures of test anxiety compared to their peers (Wren and Benson 2004) and also exhibit more test-taking behaviors indicative of anxiety (e.g., more answer-changing during test-taking). Among African American middle school children, those identified as having test anxiety had significantly poorer academic performance compared to their nonanxious African American peers, and also reported more negative self-evaluation of their cognitive and social functioning (Turner et al. 1993). A recent study by one of authors of this chapter found that African Americans report higher levels of test-related anxiety about negative performance evaluation than European Americans, and this was associated with lower neurocognitive performance (Thames et al. in press).

IQ Batteries

See Table 5.1 for a summary of measures.

Wechsler Adult Intelligence Scale—Revised and Third Editions The Wechsler Adult Intelligence Scale—revised and third editions (WAIS-R; WAIS-III; Wechsler 1997) were common assessment batteries used to assess intellectual functioning. While now supplanted by the Wechsler Adult Intelligence Scale—fourth edition (WAIS-IV; Wechsler 2008), studies on these batteries remain relevant in considering the WAIS-IV.

Early studies of the WAIS-R revealed differences between African Americans and whites on several subtests, particularly on Vocabulary, Arithmetic, and Block Design subtests (Kaufman et al. 1988; Paolo et al. 1996). Low vocabulary and arithmetic scores can be ascribed to educational differences. The Block Design subtest is of interest as it has emerged as a culturally sensitive test that negatively affects individuals who are not acculturated to the mainstream culture (Harris et al. 2003). Other performance-based nonverbal measures such as Matrix Reasoning has similarly been demonstrated to have a cultural load that is in part dependent on verbal ability (Dugbartey et al. 1999). Therefore, nonverbal performance-based subtests and indexes should not be considered culturally neutral scores.

Heaton et al. (2003) reported that compared to Hispanics and whites, African Americans had the lowest scores across all WAIS-III indexes and were three times more likely to be misclassified as impaired without demographic adjustments. Of interest, Shuttleworth-Edwards et al. (2004) compared the South African version of the WAIS-III to the USA standardization sample while stratifying for race and quality of education. Their study demonstrated that South Africans with a better educational quality produced equivalent scores to the USA sample while those with lower quality underperformed by up to 25 points. The work by Heaton and colleagues has improved the interpretation of WAIS-III performance scores in neuropsychological practice by providing demographically-corrected normative data. As with any measure, clinicians must remain mindful of the client's background when providing interpretations about performance.

Wechsler Adult Intelligence Scale—Fourth Edition The WAIS-IV (Wechsler 2008) is the most recent intelligence battery that produces a four-factor index structure comprised of the VCI, Perceptual Reasoning Index (PRI), Working Memory Index (WMI), and Processing Speed Index (PSI). The standardization sample is comprised of 2200 individuals between 16 and 90 years of age who were sampled and stratified to reflect the USA population in 2005. African Americans comprise 11.8% (n=260) of the sample. Race and education, along with sex and geographic region, were used as stratifying variables, and importantly, these variables were stratified together to ensure that equal proportions of individuals are represented within each category (i.e., African Americans with 13–15 years of education). The Advanced Clinical Solutions software allows clinicians to demographically adjust scores. However, *quality* of education remains unassessed and it is unknown the extent to which African Americans in the normative sample received equitable quality of education compared to other groups.

Table 5.1 Summary of IQ tests

Test name	Structure of test	Age range (years)	Standardization sample	Notes
Wechsler Adult Intelligence Scale—fourth edition	Four-Factor Scale assessing verbal comprehension, perceptual reasoning, working memory, and processing speed as well as a full scale IQ	16–90	2200 adults (260 African Americans)	FSIQ discrepancy of 14.5 points between the African American and whites adults. This discrepancy is less pronounced in the younger cohorts (Holdnack et al. 2013). Advanced clinical solutions provides a method to control for several demographic variables
Wechsler Intelligence Scale for Children—fourth edition	Four-Factor Scale assessing verbal comprehension, perceptual reasoning, working memory, and processing speed as well as a full scale IQ	6–16	2200 children (343 African Americans)	FSIQ discrepancy of 11.5 points between African American and Caucasian children. This discrepancy is less pronounced in younger children and also less evident with the WMI and PSI (Weiss et al. 2006)
Kaufman assessment battery for children—second edition	Based on Correct (CHC) model and has factors of crystallized intelligence, fluid reasoning, visual processing, short-term memory, and long-term retrieval	3–18	3025 children	Some evidence of minimal score discrepancies between whites and African American children (Grados and Russo-Garcia 1999; Skuy et al. 2000)
Naglieri Nonverbal Ability Test	Test relies on progressive matrices and produces a single IQ score	5–17	89,600 children	One study by the test author reported minimal racial score differences (Naglieri et al. 2000). However, subsequent research has reported that this is not an accurate conclusion as minorities with higher SES were preferentially selected (Lohman 2005)
Peabody Picture Vocabulary Test—fourth edition	Test relies on expressive oral vocabulary ability and produces a single IQ score	2–90	3540 adults and children	Research reports that African American children underperform approximately 0.5–1.0 standard deviations below the mean on the third edition (Champion et al. 2003; Washington and Craig 1999)
Reynolds Intellectual Assessment Scales	Verbal intelligence and nonverbal intelligence indexes as well as a full scale IQ	3–94	2438 adults and children	Minimal research available
Wechsler Abbreviated Scale of Intelligence—second edition	Verbal, performance, and full scale IQ scores	6–89	2300 adults and children	Short form of the WAIS-III that appears to underestimate IQ scores in minorities who are not acculturated to the USA though studies on African Americans are unavailable
Stanford–Binet Intelligence Scales—fifth edition	Based on CHC model and has factors of crystallized intelligence, fluid reasoning, visual processing, short-term memory, and quantitative knowledge	2–90+	4800 adults and children	Research on the fourth edition reported that African Americans underperform approximately one standard deviation from whites (Vincent 1991). Studies on the fifth edition are needed

Discrepancies in performance unfortunately persist and are evident by a 14.5 WAIS-IV FSIQ discrepancy between African Americans and whites in the standardization sample (white mean = 103.2; African American mean = 88.7; Lichtenberger and Kaufman 2012). Psychometric examinations report that ethnicity variance accounts for 15 % of difference between African Americans and whites 20–90 years of age, while other demographic variables including education, occupation, SES, region, and gender account for a summed total of 35.1 % of the variance (Holdnack et al. 2013). Importantly, when these variables are controlled for, ethnicity still accounted for 9.2 % of the group difference variance. This can be most likely attributed to psychosocial factors that have been discussed throughout this chapter. When stratified by age, mean score FSIQ differences between African Americans and whites decrease from 17.9 points in the 65–90 year old cohort to 9.3 points in the 16–19 year old cohort, suggesting a trend for decreasing score differences in later generations (Holdnack et al. 2013; Weiss et al. 2010). This may be related to an overall improvement in the SES and quality of education of some younger African American cohorts. In line with this, race accounts for only 1.5 % of FSIQ score differences in younger adults aged 16–19 years after controlling for other demographic variables.

Structural changes to the WAIS-IV raise some concerns. While the test developers report that efforts were made to reduce individual item bias (WAIS-IV Technical Manual), basic underlying limitations to testing cross-culturally remain apparent. For example, timed tasks that require complex verbal instructions may primarily benefit patients who familiar with Western testing standards, though research is required to substantiate this. The new Visual Puzzles and Figure Weights subtests may share similar limitations observed with the WAIS Block Design subtest in some African American populations. To our knowledge, research on the factor structure of the recently released WAIS-IV in predominantly African American populations has not yet been conducted. In addition, the WAIS-IV's stringent discontinue criteria is less forgiving for individuals who make errors unrelated to actual ability. However, these issues have not been formally examined in the literature.

Wechsler Intelligence Scale for Children—Fourth Edition African American children's FSIQ averages 91.7 points in the WISC-IV standardization sample while Hispanic children scored on average 93.1 points and whites children 103.2 points (Weiss et al. 2006). The greatest index discrepancies were with the VCI and PRI, in which African Americans scored approximately 10 points below whites. In contrast, discrepancies between scores for the WMI and PSI were approximately 5 points. These smaller gaps may in part be attributable to the increasing scores of African Americans of later generations (Weiss et al. 2010). This is supported even within the WISC-IV when children are further stratified into 6–11 and 12–16 age groups. Younger children exhibit a 6.0 score difference and the older children exhibit an 11.8 difference on the FSIQ compared to white peers. When parental education is controlled for, an 8.8 difference remains (Weiss et al. 2006).

These results indicate that African American children from the standardization sample scored lower compared to their Hispanic, Asian, and white peers. As such, similar biases related to educational and acculturation differences persist when assessing African American children. However, this gap is smaller in children and reflects a reduction of environmental adversities that may have had a greater impact on older African Americans. There is evidence that uncooperative behavior may be deleterious on IQ performance in children of minority backgrounds (Glutting and Oakland 1993). Along with factors underestimating IQ performance, there exist environmental influences that may further lower scores. African American children are at an increased risk of presenting with medical disorders (e.g., sickle cell disease), which may be undetected but yet impact cognitive performance (Steen et al. 2005). Furthermore, while the SES gap may be lessening to some extent, there remains a greater proportion of African American children from disadvantaged backgrounds that impact quality of schooling, access to healthcare, and proliferation of comorbid risk factors such as drug abuse that may hinder IQ scores.

Kaufman Batteries The Kaufman batteries are an alternative to the more widely used Wechsler batteries and include the Kaufman Assessment Battery for Children—2nd edition (KABC; Kaufman and Kaufman 2004a) and the Kaufman Brief Intelligence Test—2nd edition (K-BIT −2; Kaufman and Kaufman 2004b). These test have some preliminary evidence of invoking less cultural bias than other IQ batteries. For example, at least one published study reported equitable score profiles on the KABC-II between African American and white preschool children (Dale et al. 2011). Studies on earlier editions of the Kaufman batteries also support that these tests exhibit less bias (Grados and Russo-Garcia 1999; Skuy et al. 2000). The test authors purport that the Kaufman tests are explicitly designed to minimize the reliance on verbal abilities in determining test performance and this may in part explain these results. However, while the Kaufman batteries may be useful in some settings such as educational assessment, there are few studies available informing on their use in clinical settings. The aforementioned sources of score discrepancies relating to quality of educational attainment and SES may further impact performance. In addition, there is some evidence that the KABC is less sensitive to brain injury than the Wechsler Scales (Donders 1992). Nonetheless, these tests demonstrate a narrower gap in test scores compared to more traditional batteries and their content may warrant further investigation.

Nonverbal Tests Tests of nonverbal ability were once thought to be "culture free," and disturbingly many clinicians still practice under this assertion. Culture and cognitive style, however, also infuses nonverbal tests in several ways. Byrd et al. (2005) demonstrated that reading level influenced performance on a nonverbal neuropsychological measure of visual-spatial functioning. The significant verbal component required in administering and understanding nonverbal tests in the Wechsler batteries also argue against nonverbal tests being culture-free. For instance, the Naglieri Nonverbal Ability Test (NNAT; Naglieri 1997) is a nonverbal measure of cognitive ability that comprises progressive matrices items similar to the Wechsler Matrix Reasoning subtests. Naglieri et al. (2000) reported that African American children had a small score difference compared to white children and other ethnicities had minimal differences. However, Lohman (2005) argued that Naglieri's studies preferentially selected African American and Hispanic children from higher SES families compared to white children. Consistent with this, studies on the Peabody Picture Vocabulary Test—third edition (Dunn and Dunn 1997) report differences of approximately 0.5–1.0 standard deviations in low-income African American children compared to the standardization sample (Champion et al. 2003; Washington and Craig 1999). Such tests have appropriate uses in developmentally delayed children and/or those with limited verbal capacities, but are not culture-free.

Tests of Premorbid Intelligence Premorbid estimates of intelligence in English are traditionally assessed by reading proficiency. Consonant with previous reports detailing the differences in reading levels among African Americans exposed to a poorer quality of education than whites (Manly et al. 2002, 2004), research has demonstrated that older African Americans demonstrate lower levels of reading proficiency than whites (Boekamp et al. 1995). Tests of reading ability in individuals from disadvantaged backgrounds are very much likely to underestimate actual premorbid intelligence and should be used with caution.

Summary and Recommendations

The history of intelligence testing with African Americans quickly brings into sharp relief the complex interplay of social inequality with attempts to understand how cognition manifests in a particular cultural context. Unfortunately, this history has at times reflected the broader biased and marginalizing context in which a universalist and homogenized science has unfolded, often at the cost of the African American client. A look toward the future, however, reveals that the USA is currently experiencing a huge demographic shift in population as individuals are increasingly identifying with a diverse range of multiethnic and multiracial backgrounds. This is made particularly salient by the exponential growth of the Latino population, which, as an ethnicity, can describe diverse people of African American, white, Asian, and Indigenous backgrounds. The monolith black and white categories that historically parsed people into categories based on the color of their skin are quickly giving way to a more multicultural reality where diversity is the norm rather than the exception. Over time, greater intercultural penetration and new cultural practices brought on by advances in technology will emerge. As this occurs, reliance on self-reported ethnic and/or racial categories will need to shift toward identifying measurable dispositional factors that can be systematically studied across historical racial and ethnic classifications. In response to this, psychologists must increasingly be mindful, aware, and responsive to the limitations associated with using self-reported ethnicity and/or race as proxy measures and instead critically examine potential cultural practices when working with historically underrepresented individuals. Indeed, it is our professional responsibility to not only acknowledge and understand these issues, but to take pragmatic steps to improve our field and overcome the obstacles related to reliable and meaningful cognitive assessment.

This chapter reviewed some of the challenges specific to IQ testing and African Americans. Score discrepancies between African Americans and whites, nevertheless persist in currently available assessment modalities, and are more often than not linked to psychosocial factors such as SES, quality of education, cultural factors, and biases of the tests used. It is quite significant and encouraging that evidence is accumulating to suggest that score discrepancies are narrowing in younger cohorts, which in part can be attested to by better access to resources and educational opportunities for an increasing proportion of African American youth. However, several children and older individuals remain at a disadvantage when tested, and this issue needs to be addressed head-on to further progress in the field of intellectual assessment with African Americans and our diverse population more broadly.

The following closing summary points are highlighted in the hope of both bringing attention to specific key factors in the intellectual assessment of African Americans, and also to spark further discussion and research questions that are still needed to create a more inclusive and culturally representative empirical base:

1. While the racial IQ gap still persists, this gap may be shrinking as African Americans continue to gain improved access to education and improved socioeconomic opportunities.
2. The racial gaps of WAIS-IV and WISC-IV scores are reportedly narrower in younger cohorts, suggesting that younger African Americans may be facing fewer of the environmental stressors than their parents and grandparents; however, much more research is needed to gauge the stability of this trend over time.
3. Quality, not years, of education appears to be a more valid demographic characteristic assess when administering IQ tests to African Americans. Currently, literacy measures appear to be the most reliable and empirically validated proxy for educational quality.
4. Psychologists should be mindful of possible dispositional influences that can impact test performance (e.g., stereotype threat, performance anxiety).
5. Nonverbal tests are not necessarily exempt from the cultural biases which deceptively may seem more face valid and explicit on verbal tests.
6. In a neuropsychological setting, demographic adjustments can potentially reduce the score gap between African Americans and non-Hispanic whites. Such adjustments minimize the likelihood of misclassifying an individual as cognitively compromised, but may simultaneously increase the likelihood of missing a true cognitive weakness if improperly used.
7. Demographic corrections may not be appropriate in settings where identifying cognitive impairment is not the purpose of the evaluation, and may actually lead to misrepresentation of functional abilities relative to the general population.

8. Further research is required to determine if the new WAIS-IV subtests demonstrate biased results against historically underrepresented minorities such as African Americans. Research may particularly benefit from replication of early validation studies on the Kaufman batteries.
9. Psychologists should be aware that premorbid verbal intelligence and quality of education are both assessed by standardized reading measures, which are not interchangeable and are actually psychometrically calibrated to measure different constructs (i.e., estimated verbal intellectual functioning versus grade equivalence) despite sharing a common methodology (i.e., word-reading).
10. More research is needed to unpack proxy variables historically categorized as race and ethnicity into measurable cultural practices that can be empirically and systematically studied with the degree of scientific rigor worthy of any other explanatory variable.

References

Affeldt, J. (2000). Legal issues in testing African Americans. In W. A. Thomas (Ed.), Larry P. revisited: IQ testing of Americans (pp. 23–27). San Francisco: California Publishing Co.

Allen, D. N., Thaler, N. S., Barchard, K. A., Vertinski, M., & Mayfield, J. (2012). Factor structure of the comprehensive trail making test in children and adolescents with brain dysfunction. *Psychological assessment, 24*(4), 964.

Anglin, D. M., & Kwate, N. O. A. (2009). Social and cultural factors in the cognitive and clinical assessment of African Americans. In H. A. Neville, B. M. Tynes & S. O. Utsey (Eds.), *Handbook of African American psychology* (pp. 459–468). Thousand Oaks: Sage.

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology, 41*(2), 174–181.

Boekamp, J. R., Strauss, M. E., & Adams, N. (1995). Estimating premorbid intelligence in African-American and white elderly veterans using the American version of the national adult reading test. *Journal of Clinical and Experimental Neuropsychology, 17*(5), 645–653.

Brazziel, W. F. (2000). Improving SAT scores: Pros, cons, methods. In W. A. Thomas (Ed.), Larry P. revisited: IQ testing of African Americans (pp. 60–72). San Francisco: California Publishing Co.

Byrd, D. A., Touradji, P. E. G. A. H., Tang, M. X., & Manly, J. J. (2004). Cancellation test performance in African American, Hispanic, and White elderly. *Journal of the International Neuropsychological Society, 10*(3), 401–411.

Byrd, D. A., Jacobs, D. M., Hilton, H. J., Stern, Y., & Manly, J. J. (2005). Sources of errors on visuoperceptual tasks: Role of education, literacy, and search strategy. *Brain and cognition, 58*(3), 251–257.

Cagigas, X. E., & Manly, J. J. (2014). Cultural Neuropsychology: the New Norm. In M. W. Parsons & T. A. Hammeke (Eds.), Clinical Neuropsychology: A Pocket Handbook for Assessment, 3rd Edition. Washington D.C.: APA.

Champion, T. B., Hyter, Y. D., McCabe, A., & Bland-Stewart, L. M. (2003). "A matter of vocabulary" performances of low-income African American head start children on the peabody picture vocabulary test-III. *Communication Disorders Quarterly, 24*(3), 121–127.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*(5791), 1307–1310.

Coleman, J. (1966). *Equality of educational opportunity*. Washington DC: Government Printing Office.

Dale, B. A., McIntosh, D. E., Rothlisberg, B. A., Ward, K. E., & Bradley, M. (2011). Profile analysis of the Kaufman assessment battery for children, second edition, with African American and Caucasian preschool children. *Psychology In The Schools, 48*(5), 476–487. doi:10.1002/pits.20571.

Davis, C., Aronson, J., & Salinas, M. (2006). Shades of threat: Racial identity as a moderator of stereotype threat. *Journal of Black Psychology, 32*(4), 399–417.

Delis, D. C., Jacobson, M., Bondi, M. W., Hamilton, J. M., & Salmon, D. P. (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: Lessons from memory assessment. *Journal of the International Neuropsychological Society, 9*(06), 936–946.

Diehr, M. C., Heaton, R. K., Miller, W., & Grant, I. (1998). The paced auditory serial addition task (PASAT): Norms for age, education, and ethnicity. *Assessment, 5*(4), 375–387.

Donders, J. (1992). Validity of the Kaufman assessment battery for children when employed with children with traumatic brain injury. *Journal of clinical psychology, 48*(2), 225–230.

Dotson, V. M., Kitner-Triolo, M., Evans, M. K., & Zonderman, A. B. (2008). Literacy-based normative data for low socioeconomic status African Americans. *The Clinical Neuropsychologist, 22*(6), 989–1017.

Dugbartey, A. T., Sanchez, P. N., Gail Rosenbaum, J., Mahurin, R. K., Mark Davis, J., & Townes, B. D. (1999). WAIS-III matrix reasoning test performance in a mixed clinical sample. *The Clinical Neuropsychologist, 13*(4), 396–404.

Dunn, L. M., & Dunn, L. M. (1997). Examiner's manual for the PPVT-III peabody picture vocabulary test: Form IIIA and Form IIIB. AGS.

García, L. F., & Abad, F. J. (2003). Factor structure of the Spanish WAIS-III. *Psicothema, 15*(1), 155–160.

Glutting, J., & Oakland, T. (1993). *Guide to the assessment of test session behavior*. San Antonio: Psychological Corporation.

Golay, P., & Lecerf, T. (2011). Orthogonal higher order structure and confirmatory factor analysis of the French Wechsler Adult Intelligence Scale (WAIS-III). *Psychological assessment, 23*(1), 143.

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why?. *Current Directions in Psychological Science, 13*(1), 1–4.

Grados, J. J., & Russo-Garcia, K. A. (1999). Comparison of the Kaufman Brief Intelligence Test and the Wechsler Intelligence Scale for children-third edition in economically disadvantaged African American youth. *Journal of clinical psychology, 55*(9), 1063–1071.

Harris, J. G., Tulsky, D. S., & Schultheis, M. T. (2003). Assessment of the non-native English speaker: Assimilating history and research findings to guide clinical practice. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 343–390). San Diego: Academic. doi:10.1016/B978-012703570-3/50015-8.

Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Demographic effects and use of demographically corrected norms with the WAIS-III and WMS–III.

Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). An exchange: Part I: Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational researcher, 23*(3), 5–14.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press

Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. The Journals of Gerontology Series B: *Psychological Sciences and Social Sciences, 58*(1), P3–P11.

Hilliard, A. G. (1997). Language, culture and the assessment of African American children. In A. L. Goodwin & Maritza B. MacDonald (Eds.), Assessment for equity and inclusion: Embracing all our children (pp. 229–240). Routledge.

Holdnack, J. A., Drozdick, L., Weiss, L. G., & Iverson, G. L. (Eds.). (2013). *WAIS-IV, WMS-IV, and ACS: Advanced clinical interpretation*. Access Online via Elsevier.

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of personality and social psychology, 93*(4), 544.

Kaufman, A. S., & Kaufman, N. L. (2004a). *Kaufman assessment battery for children* (2nd ed.). Circle Pines: AGS Publishing.

Kaufman, A. S., & Kaufman, N. L. (2004b). *Kaufman Brief Intelligence Test* (2nd ed.). Circle Pines: AGS Publishing.

Kaufman, A. S., McLean, J. E., & Reynolds, C. R. (1988). Sex, race, residence, region, and education differences on the 11 WAIS-R subtests. *Journal of clinical psychology*. doi:10.1002/1097-4679(198803).

Kit, K. A., Mateer, C. A., Tuokko, H. A., & Spencer-Rodgers, J. (2014). Influence of negative stereotypes and beliefs on neuropsychological test performance in a traumatic brain injury population. *Journal of the International Neuropsychological Society, 20*(2), 157–167.

Lichtenberger, E. O., & Kaufman, A. S. (2012). *Essentials of WAIS-IV assessment* (Vol. 96). New York: Wiley.

Lohman, D. F. (2005). Review of Naglieri and Ford (2003): Does the Naglieri nonverbal ability test identify equal proportions of high-scoring White, Black, and Hispanic students? *Gifted Child Quarterly, 49*(1), 19–28.

Manly, J. J. (2005). Advantages and disadvantages of separate norms for African Americans. *The Clinical Neuropsychologist, 19*(2), 270–275.

Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., & Stern, Y. (1998). Cognitive test performance among nondemented elderly African Americans and whites. *Neurology, 50*(5), 1238–1245.

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society, 8*(3), 341–348.

Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology, 11*(1), 37–46.

Marlowe, W. B. (2000). Multicultural perspectives on the neuropsychological assessment of children and adolescents. In Handbook of cross-cultural neuropsychology (pp. 145–165). Springer US.

McFarland, L. A., Lev-Arey, D. M., & Ziegert, J. C. (2003). An examination of stereotype threat in a motivational context. *Human Performance, 16*(3), 181–205.

Naglieri, J. A. (1997). Naglieri nonverbal ability test. San Antonio, TX: The Psychological Corporation.

Naglieri, J. A., & Ronning, M. E. (2000). Comparison of White, African American, Hispanic, and Asian children on the Naglieri Nonverbal Ability Test. *Psychological Assessment, 12*(3), 328–334.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. *American psychologist, 67*(2), 130.

Norman, M. A., Moore, D. J., Taylor, M., Franklin, D. Jr., Cysique, L., Ake, C., et al. (2011). Demographically corrected norms for African Americans and Caucasians on the Hopkins verbal learning test-revised, brief visuospatial memory test-revised, Stroop color and word test, and Wisconsin card sorting test 64-card version. *Journal of clinical and experimental neuropsychology, 33*(7), 793–804.

O'Neill, J. (1990). The role of human capital in earnings differences between black and white men. *The Journal of Economic Perspectives, 4*(4), 25–45.

Paolo, A. M., Ryan, J. J., Ward, L. C., & Hilmer, C. D. (1996). Different WAIS-R short forms and their relation to ethnicity. *Personality and individual differences, 21*(6), 851–856.

Pedraza, O., Graff-Radford, N. R., Smith, G. E., Ivnik, R. J., Willis, F. B., Petersen, R. C., & Lucas, J. A. (2009). Differential item functioning of the Boston naming test in cognitively normal African American and Caucasian older adults. *Journal of the International Neuropsychological Society, 15*(05), 758–768.

Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reactions research. *Human Performance, 16*(3), 231–259.

Reitan, R. M., & Wolfson, D. (2005). The effect of age and education transformations on neuropsychological test scores of persons with diffuse or bilateral brain damage. *Applied neuropsychology, 12*(4), 181–189.

Reynolds, C. R. (2000). Methods for detecting and evaluating cultural bias in neuropsychological tests. In E. Fletcher-Janzen, T. L. Strickland, & C. Reynolds (Eds.), Handbook of Cross-Cultural Neuropsychology (pp. 249–285). Springer Science & Business Media.

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G., Hartman, N. P., & Radloff, S. E. (2004). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAIS-III test performance. *Journal of Clinical and Experimental Neuropsychology, 26*(7), 903–920.

Skuy, M., Taylor, M., O'Caroll, S., Fridjohn, P., & Rosenthal, L. (2000). Performance of black and white South African children on the Wechsler Intelligence Scale for children-revised and the Kaufman assessment battery. *Psychological Reports, 86*(3), 727–737.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology, 35*(1), 4–28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist, 52*(6), 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

Steen, R. G., Fineberg-Buchner, C., Hankins, G., Weiss, L., Prifitera, A., & Mulhern, R. K. (2005). Cognitive deficits in children with sickle cell disease. *Journal of child neurology, 20*(2), 102–107.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology, 77*(6), 1213.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Boston: Houghton Mifflin Co.

Thames, A. D., Panos, S. E., Arentoft, A., Byrd, D. A., Hinkin, C. H., & Arbid, N. (in press). Mild test anxiety influences neurocognitive performance among African Americans and European Americans: Identifying interfering and facilitating sources.

Thames, A. D., Hinkin, C. H., Byrd, D. A., Bilder, R. M., Duff, K. J., Mindt, M. R., et al. (2013). Effects of stereotype threat, perceived discrimination, and examiner race on neuropsychological performance: Simple as Black and White? *Journal of the International Neuropsychological Society, 19,* 1–11.

Thomsen, C., Gallup, L., & Llorente, A. M. (2008). Intellectual abilities: Theoretical and applied assessment considerations. In A. M. Llorente & A. M. Llorente (Eds.), *Principles of neuropsychological assessment with Hispanics: Theoretical foundations and clinical practice* (pp. 57–77). New York: Springer. doi:10.1007/978-0-387-71758-6_4

Turner, B. G., Beidel, D. C., Hughes, S., & Turner, M. W. (1993). Text anxiety in African American school children. School Psychology Quarterly, 8(2), 140.

Turkheimer, E., Haley, A., Waldron, M., d'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological science, 14*(6), 623–628.

Vincent, K. R. (1991). Black/White IQ differences: Does age make the difference? *Journal of Clinical Psychology, 47*(2), 266–270.

Washington, J. A., & Craig, H. K. (1999). Performances of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, and Hearing Services in Schools, 30*(1), 75–82.

Wechsler, D. (1997). Wechsler adult intelligence scale–Third edition. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth edition. San Antonio, TX: Pearson.

Weiss, L. G., Saklofske, D. H., Prifitera, A., & Holdnack, J. A. (2006). *WISC-IV advanced clinical interpretation*. Access Online via Elsevier.

Weiss, L. G., Saklofske, D. H., Coalson, D., & Raiford, S. E. (Eds.). (2010).*WAIS-IV clinical use and interpretation: Scientist-practitioner perspectives*. Access Online via Elsevier.

Williams, R. L. (1972). *The BITCH—100: A culture-specific test*. American Psychological Association, Annual Convention. Honolulu, Hawaii.

Wren, D. G., & Benson, J. (2004). Measuring test anxiety in children: Scale development and internal construct validation. *Anxiety, Stress & Coping, 17*(3), 227–240.

6

Assessing Effort and Malingering with the African American Client

Lorraine T. Benuto, Brian D. Leany and Andrew Lee

Malingering defined as "…the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives…" (American Psychiatric Association 2000). In the DSM-V (American Psychiatric Association 2013) malingering remains a V Code (as opposed to a diagnostic category) and its definition remains largely unchanged from the DSM-IV.

As discussed previously by Benuto and Leany (2013), the assessment of malingering is important as malingering can cause harm to the individual who is malingering (e.g., inappropriate application of treatment, delay in the administration of actual treatment needed, etc.) and the costs associated with malingering are exponentially high. Moreover, there is some debate regarding the use of the term "effort" over "malingering" in part due to the negative connotation of the latter and because the clinician can never really know the motivation behind feigning symptoms. Nonetheless, the American Psychiatric Association (2000) has made clear that the clinician can suspect deliberate feigning when any of the following are apparent: medicolegal context of presentation, marked discrepancy between self-reported stress or disability and objective findings, lack of cooperation during diagnostic evaluation or with prescribed treatment, or presence of antisocial personality disorder. Alternatively the term "symptom validity" is often used in the literature and accurately captures the construct of malingering as the question of a "symptom validity measure" is (as implied by the name): are the symptoms the client is presenting with valid?

The aim of this chapter is to explore the most commonly used and/or researched measures that assess for malingering (or symptom validity) specifically in the context of their use with African Americans and to generate evidence-based recommendations for the assessment of malingering with the African American client.

The State of the Field

Admittedly, the focus on ethnic or cultural differences in the field of clinical psychology is not new. Nonetheless, progress in this field has been slow and this subfield (assessment of malingering/symptom validity) is no exception. In fact, very little relevant research on the assessment of malingering with African Americans was identified. Under ethical considerations in

---

L. T. Benuto (✉)
Administrative Faculty, Department of Psychology
MS0296 University of Nevada, Reno, NV 89557, USA
e-mail: dr.benuto@gmail.com

L. T. Benuto · B. D. Leany
Northcentral University, Core Faculty in the Graduate School, 10000 E University Drive, Prescott Valley, AZ 86314, USA

B. D. Leany
Lake's Crossing Center, Maximum Security Facility for Forensic Mental Health Services, 500 Galletti Way, Sparks, NV 89504, USA
e-mail: leanyb@gmail.com

A. Lee
Department of Psychology, University of Nevada, Reno
2500 Dickerson rd. # 160, Reno, NV 89503, USA
e-mail: hwmail2@yahoo.com

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_6,

Table 6.1 Symptom validity measures

Test	Administration	Scoring/Interpretation	Cultural considerations
Test of memory malingering (TOMM: Tombaugh 1996)	Consists of two learning trials and an optional retention trial. The client is presented with 50 pictures of common objects and then presented with two pictures and asked to identify which of the two he or she previously viewed	Results are based on two cutoff scores: (1) below chance and (2) criterion based on head injured and cognitively impaired clients	The TOMM is recommended as a reliable and valid SVT for use with African Americans
Minnesota Multiphasic Personality Inventory-2 (MMPI-2)	Self-report questionnaire	Elevated scores on the fake bad scale; infrequence scale; infrequency minus defensiveness; back infrequency scale; and the revised social desirability scale	African Americans may be more likely to present with an invalid profile. Additional SVT measures should be administered
Reliable digit span (Greiffenstein et al. 1994)	Numbers are verbally given to the examinee who repeats the numbers back to the examiner either forward and/or backward	The RDS is calculated by summing the longest forward and backward digit strings and a score of ≤7 is indicative of feigning	When using the digit span from the WAIS-III a cutoff of ≤ 6 is recommended
Rey 15 items (Salazar et al. 2007)	Consists of two parts: first, the client is exposed to a page that contains 15 items (see Fig. 6.1) and asked to reproduce items. In a recognition trial, the client is presented with a page that contains all target items plus 15 new items and asked to identify which of the items were not presented earlier	A combined recall and recognition score (free recall + [recognition—false positives] < 20) results in good sensitivity (71 %) and high specificity (≥ 92 %) (Boone et al. 2002)	For African Americans a cutoff of < 18 is recommended

cross-cultural neuropsychology, it was noted that African Americans perform poorer on tasks of visual confrontation naming and on tasks of nonverbal abilities compared to whites (Brickman et al. 2006). Given that measures such as the test of memory and malingering (TOMM: Tombaugh 1996) require a forced choice, visual confrontation recall, we have potential concern that a lack of awareness for such racial differences in performance, may lead to misclassification of malingered versus genuine symptoms. Tables 6.1 and 6.2 include a description of the measures that are reviewed in this chapter with associated recommendations regarding the use of these measures with African Americans. These measures do not represent an exhaustive presentation of the measures that can be used to assess for feigning of symptoms but rather reflect commonly used measures for which conclusions could be drawn regarding their use with African Americans. Literature on other commonly used assessments that can detect feigning were also reviewed (e.g., the Word Memory Test: Green 2003; Green and Astner 1995; Medical Symptom Validity Test: Green 2004), however, we opted not to include a discussion on measures where substantial literature to develop meaningful recommendations was absent.

Test of Memory Malingering (TOMM: Tombaugh 1996)

The TOMM assesses for false symptoms of memory impairment. Researchers have found that over half (55.6 %) of participants in nonclinical

Table 6.2 Symptom validity interviews

Test	Administration	Scoring	Cultural considerations
The structured interview of reported symptoms (SIRS; Rogers et al. 1992)	Consists of 172-items administered in a structured interview format that usually lasts 45 min	Questions are mostly scored using a "no," "qualified yes" (e.g., "sometimes"), or "yes" format and 32 items are asked twice as a means to assess response consistency	No adjustments necessary
Miller-forensic assessment of symptoms test (M-FAST; Miller 2001)	Consists of a 25-item structured interview intended to help identify malingered psychopathology, especially in illiterate populations	A total score corresponds with seven strategies frequently used among individuals who malinger	No adjustments necessary

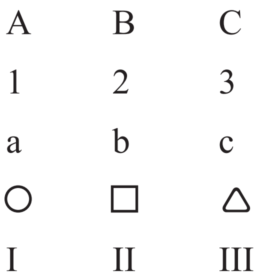

Fig. 6.1 Rey 15 items (Taken from Benuto and Leany 2013)

populations score 49 or 50 on the TOMM and only 8.4 % score lower than 45 (Tombaugh 1996); from a purely chance perspective at worst one would expect the client to score 50 % on the TOMM although even individuals with actual cognitive impairment tend to score higher than chance. Although the TOMM has not been specifically researched with African American's, a number of research studies examining the reliability and validity of the TOMM have included large percentages of African American participants (although admittedly other studies have included few African American participants in their samples). For example, Weinborn et al. (2003) attempted to validate the TOMM with a forensic sample ($N=61$) of which over half (52 %) was African American. While Weinborn et al. did not examine ethnic differences in performance on the TOMM, they concluded that the TOMM has excellent specificity and modest sensitivity. From this study we can conclude that these findings generally support the validity of the TOMM with a forensic psychiatric population of which half were African Americans. More recently Wisdom et al. (2012) examined the TOMM with a sample of epilepsy patients ($N=213$) and a portion ($n=46$; 22 %) of this sample was African American. The authors reported no significant differences between age, gender, or ethnic groups on any of the TOMM trials administered suggesting that this is a valid and reliable measure for use with African Americans.

Minnesota Multiphasic Personality Inventory-2 (MMPI-2)

Some researchers have found the MMPI-2 to be systematically biased in determining pathologies in African Americans (Monnot et al. 2009), or more precisely in somatic symptoms (Zang and Snowden 1999; Sellbom et al. 2010). Yet in regards to malingering, research has shown that this is not the case (Dean et al. 2008). In an impressive study that included 229 African American psychiatric inpatients (alongside 1558 Caucasian psychiatric inpatients) Arbisi et al. (2002) found that while mean differences on several MMPI-2 validity and clinical scales were present and some bias in the scales was evident (reflecting minor underprediction of psychopathology in African Americans) the authors noted that in almost all cases these differences were small and not clinically significant. Despite the findings described above, McNulty et al. (2003) found that a substantially greater proportion of African

Americans generated an invalid MMPI-2 profile when compared to their Caucasian counterparts in a sample of state corrections sample ($N=51{,}486$; 60.8% African American). The authors suggested that content nonresponsiveness (CNR: an inability or lack of willingness to respond to the actual content off the items) and a difference in motivation could most likely explain the observed discrepancy in CNR between Caucasians and African Americans.

From the studies discussed above, we can deduce that on the MMPI-2 African Americans may be more likely to present with an invalid profile. However, the invalid nature of the profile may be attributable to CNR or test bias. Thus an invalid MMPI-2 should not be used exclusively to determine feigning of symptoms. Additional data should be collected to either support or refute the findings from an invalid or exaggerated MMPI-2 profile.

Reliable Digit Span (RDS)

Digit Span (DS) is used in many psychological tests designed to assess various aspects of cognitive function (e.g., Wechsler Adult Intelligence Scale-IV; Montreal Cognitive Assessment) and specifically assesses auditory attention. The RDS is calculated by summing the longest forward and backward digit strings from the digit span test. Because digit span performance tends to be preserved in persons with brain dysfunction (e.g., Greiffenstein et al. 1994; Heinly et al. 2005; Iverson and Tulsky 2003), it serves as good indicator of feigned cognitive impairment and there is substantial research to support its validity (Suhr and Barrash 2007). RDS (RDS; Greiffenstein et al. 1994) can be used to assess for effort or feigning of cognitive or other symptoms and a cutoff score of less than or equal to 7 can accurately classify 75% of persons who are not putting forth optimal effort (Axelrod et al. 2006). Some have recommended that a lower cutoff than 7 might be needed when calculated using the new WAIS-IV (Reese et al. 2012). While, there is evidence that suggests that African Americans may score (1.5 scaled score points) lower on digit span than Caucasians (Boone et al. 2007), this may not necessary translate to lower performance on RDS. However, although there may be a correlation between low scores on digit span and RDS, it likely would not be very strong and it would be hard to extrapolate one from the other. Consider the following, if group A had a raw score of 6 but got each trial right, their reliable digit span would be higher than group B with a raw score of 8 but several missed trials. This is because the RDS delineates where the individual began making errors (e.g., one of the pair of two for the trial) rather than the more traditional scoring system of one point (raw score) for at least one correct pair. Nonetheless, recommendations for an RDS adjustment for African Americans do exist. Salazar et al. (2007) recommend that for African Americans a RDS ≤ 6 cutoff is appropriate (presumably for the old WAIS-III).

Rey 15 Items

The Rey 15 items consists of two parts: first, the client is exposed to a page that contains 15 items (see Fig. 6.1) and then asked to reproduce items. In a recognition trial, the client is presented with a page that contains all target items plus 15 new items and asked to identify which of the items were not presented earlier. While a cutoff of > 20 is typically used, Salazar et al. (2007) recommend that for African Americans the cutoff on the Rey-15 should be moved to < 18.

The Structured Interview of Reported Symptoms (SIRS; Rogers et al. 1992)

As previously discussed by Benuto and Leany (2013) the original SIRS is considered the "gold standard" for assessing for feigned mental disorders (Rogers 2008) and as such is commonly used in forensic practice (Archer et al. 2006). Currently the SIRS is in its second edition (*SIRS, Second Edition: SIRS-2*; Rogers et al. 2010a). Revisions to the SIRS were made to prevent false-positive and false-negative classification errors (Rubenzer 2010). The SIRS consists of 172-items (of which 32 are repeated inquiries to help the clinician to detect inconsistency of responding); is administered via a structured interview (questions are mostly scored using a “no,” “qualified yes,” e.g., “sometimes”, or “yes”); and takes approximately 45 min to administer (Rogers et al. 2010a, b). Within the SIRS there are eight primary scales:

- Rare symptoms (RS)
- Symptom combinations (SC)
- Improbable and absurd symptoms (IA)
- Blatant symptoms (BL)
- Subtle symptoms (SU)
- Selectivity of symptoms (SEL)
- Severity of symptoms (SEV)
- Reported versus observed symptoms (RO)

The SIRS-2 has an interpretive decision model that assists clinicians in classifying response styles (Kocsis 2011). Several researchers have examined the generalizability of the SIRS with minorities and no significant differences have been reported with regard to race (Rogers 2001), suggesting that the SIRS may be used with African American clients.

Miller-Forensic Assessment of Symptoms Test (M-FAST; Miller 2001)

As discussed by Benuto and Leany (2013) the M-FAST (Miller 2001) is a 25-item structured interview and can be used to identify malingered psychopathology, especially in illiterate populations (Miller 2005). A total score corresponds with seven strategies frequently used among individuals who malinger:

- Unusual hallucinations
- Reported versus observed
- Rare combinations
- Extreme symptomatology
- Negative image
- Unusual symptom course
- Suggestibility

The M-FAST has good psychometric properties (Miller 2001, 2004) and there is evidence that the M-FAST is generalizable across racial groups (Miller 2005). Specifically with regard to African American’s, Guy and Miller administered the M-Fast to a sample ($N=50$) of incarcerated males of which 36% ($n=18$) were African American. Similar performance was noted across ethnic groups suggesting that current norms are likely generalizable across the ethnic groups studied (Caucasians, African Americans, and Hispanics). Moreover the M-Fast has been studied in samples that have included African Americans (Vitacco et al. 2008) and even though these researchers did not look specifically at ethnic differences these studies have focused on the psychometric properties of this measure and revealed sufficient validity and reliability on the M-Fast.

Summary and Recommendations

In this chapter we reviewed some of the most commonly used and well-researched tests that aim to assess for malingering and/or effort. A summary of these measures with appropriate recommendations can be found in Tables 6.1 and 6.2. This review has shown that there are many very useful measures that can be used in determining the feigning African American client. When used correctly with the considerations described above, these measures can help distinguish malingerers from those with real psychiatric symptoms. Thus the inclusion of these measures in an assessment battery, particularly when motivation to feign is discovered, should aid in proper diagnosis and corresponding interventions for the African American client.

While in general we advocate that, when administering a test that requires the results to be interpreted with caution, additional data should be gathered, there is an important consideration that should be made when administering symptom validity measures. Specifically Boone (2009) and Victor et al. (2009) have noted that when using several measures of effort, clinicians should avoid using highly correlated SVTs as using highly correlated SVTs would provide limited incremental information. Thus, a multidimensional approach is recommended to minimize the possibility of false positives or negatives. Specifically, the clinician should administer multiple tests that tap into different dimensions of malingering, obtain any history of malingering behaviors, and corroborate the information provided with reliable collateral contacts.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: APA.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: APA.

Arbisi, P. A., Ben-Porath, Y. I., & McNulty, J. (2002). A comparison of MMPI-2 validity in African American and Caucasian psychiatric inpatients. *Psychological Assessment, 14*(1), 3–15.

Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87*(1), 84–94, 11.

Axelrod, B. N., Fichtenberg, N. L., Millis, S. R., & Wertheimer, J. C. (2006). Detecting incomplete effort with digit span from the Wechsler Adult Intelligence Scale-third edition. *The Clinical Neuropsychologist, 3*(3), 513–523.

Benuto, L. T., & Leany, B. D. (2013). Assessing effort and malingering with the hispanic client. In *Guide to Psychological Assessment with Hispanics* (pp. 99–112). Springer US.

Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological evaluation. *The Clinical Neuropsychologist, 23,* 729–741.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-item recognition trial: A technique to enhance sensitivity of the Rey 15-item memorization test. *Journal of clinical and Experimental Neuropsychology, 24*(5), 561–573.

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology, 3*(3), 355–365.

Brickman, A. M., Cabo, R., & Manly, J. J. (2006). Ethical issues in cross-cultural neuropsychology. *Applied Neuropsychology, 13*(2), 91–100.

Dean, A. C., Boone, K. B., Kim, M. S., Curiel, A. R., Martin, D. J., Victor, T. L., Zeller, M. A., & Lang, Y. K. (2008). Examination of the impact of ethnicity on the Minnesota Multiple Personality Inventory-2 (MMPI-2) Fake Bad Scale. *The Clinical Neuropsychologist, 22,* 1054–1060.

Green, P. (2003). *Green's Word Memory Test for Windows: User's manual*. Edmonton: Green's Publishing.

Green, P. (2004). *Green's Medical Symptom Validity Test (MSVT): User's manual*. Edmonton: Green's Publishing.

Green, P., & Astner, K. (1995). *Manual for the Oral World Memory Test*. Edmonton: Neurobehavioral Associates.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 3*(3), 218–224.

Heinly, M. T., Greve, K. W., Bianchini, K. J., Love, J. L., & Brennan, A. (2005). WAIS Digit Span-based indicators of malingered neurocognitive dysfunction: Classification accuracy in traumatic brain injury. *Assessment, 12,* 429–444.

Iverson, G. L., & Tulsky, D. S. (2003). Detecting malingering on the WAIS-III: Unusual Digit Span performance patterns in the normal population and in clinical groups. *Archives of Clinical Neuropsychology, 18*(1), 1–9.

James, D. J., & Glaze, L. E. (2006). Mental health problems in prison and jail inmates. Bureau of Justice Statistics Special Report, September.

Kocsis, R. N. (2011). The structured interview of reported symptoms 2nd edition (SIRS-2): The new benchmark towards the assessment of malingering. *Journal Of Forensic Psychology Practice, 11*(1), 73–81.

McNulty, J., Forbey, J., Graham, J., Ben-Porath, Y., Black, M., & Anderson, S. (2003). MMPI-2 validity scale characteristics in a correctional sample. *Assessment, 10,* 288–298.

Miller, H. A. (2001). MFAST: Miller Forensic Assessment of Symptoms Test professional manual. Odessa: Psychological Assessment Resources, Inc.

Miller, H. A. (2004). Examining the use of the M-FAST with criminal defendants incompetent to stand trial. *International Journal of Offender Therapy and Comparative Criminology, 48*(3), 268–280.

Miller, H. A. (2005). The Miller Forensic Assessment of Symptoms Test (M-FAST): Test generalizability and utility across race, literacy, and clinical opinion. *Criminal Justice and Behavior, 32*(6), 591–611.

Monnot, M. J., Quirk, S. W., Hoerger, M., & Brewer, L. (2009). Racial bias in personality assessment: Using the MMPI-2 to predict psychiatric diagnoses of African American and Caucasian chemical dependency inpatients. *Psychological Assessment, 21*(2), 137–151.

Reese, C., Suhr, J. A., & Riddle, T. (2012). Exploration of malingering indices in the Weschler Adult Intelligence Scale-fourth edition Digit Span subtest. Archives of Clinical Neuropsychology Advanced Access.

Rogers, R. (2001). *Handbook of diagnostic and structured interviewing*. New York: Guilford.

Rogers, R. (Eds.). (2008). *Clinical assessment of malingering and deception*. New York: Guilford.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured interview of reported symptoms (SIRS) and professional manual*. Odessa: Psychological Assessment Resources.

Rogers, R., Sewell, K. W., & Gillard, N. (2010a) *Structured interview of reported symptoms-2 (SIRS-2) and professional manual*. Odessa: Psychological Assessment Resources.

Rogers, R., Sewell, K. W., & Gillard, N. D. (2010b). *SIRS professional manual* (2nd ed.). Lutz: Psychological Assessment Resources.Inc

Rubenzer, S. (2010). Review of the structured inventory of reported symptoms-2 (SIRS-2). *Open Access Journal of Forensic Psychology, 2,* 273–286.

Salazar, X. F., Lu, P. H., Wen, J., & Boone, K. (2007). The use of effort tests in ethnic minorities and in non-English-speaking and English as a second language populations. In K. Boone & K. Boone (Eds.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 405–427). New York: Guilford.

Sellbom, M., Toomey, J. A., Wygant, D. B., Kucharski, L. T., & Duncan, S. (2010). Utility of the MMPI-2-RF (Restructured Form) validity scales in detecting malingering in a criminal forensic setting: A known-groups design. *Psychological Assessment, 22*(1), 22–31.

Suhr, J. A., & Barrash, J. (2007). Performance on standard attention, memory, and psychomotor speed tasks as indicators of malingering. In G. L. Larrabee (Ed.), *Assessment of malingered neurospychological deficits* (pp. 131–170). New York: Oxford University Press.

Tombaugh, T. N. (1996). *The test of memory malingering (TOMM)*. Toronto: Multi-Health Systems.

Victor, T. L., Boone, K. B., Serpa, J. G., Beuhler, J., & Ziegler, E. A. (2009). Interpreting the meaning of multiple symptoms validity tests failure. *The Clinical Neuropsychologist, 23,* 297–313.

Vitacco, M. J., Jackson, R. L., Rogers, R., Neumann, C. S., Miller, H. A., & Gabel, J. (2008). Detection strategies for malingering with the miller forensic assessment of symptoms test a confirmatory factor analysis of its underlying dimensions. *Assessment, 15*(1), 97–103.

Wechsler, D. (2009). *Wechsler Memory Scale-fourth edition*. San Antonio: Pearson.

Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, F. (2003). A validation of the test of memory malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology, 3*(7), 979–990.

Wisdom, J. P., Cavaleri, M. A., Onwuegbuzie, A. J., & Green, C. A. (2012). Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. *Health Services Research, 3*(2), 721–745.

Zhang, A. Y., & Snowden, L. R. (1999). Ethnic characteristics of mental disorders in five U.S. communities. *Cultural Diversity and Ethnic Minority Psychology, 5,* 134–146.

7

Assessing Mood Disorders and Suicidality in African Americans

Alfiee M. Breland-Noble, Jason Sotomayor
and F. Antoinette Burriss

Assessing Mood Disorders and Suicidality in African Americans

According to the United States (US) Census data, African Americans constitute approximately 13 % of the US population, including people who are the direct descendants of those forcibly transported to the USA and forced into slavery including Caribbean and African immigrants (U.S. Census Bureau 2010). Sometimes referred to in aggregate as the Black population, persons of African descent in the USA, are projected to double the size of the US African descended population by 2050 (Joe 2006). As a result, it is imperative for clinicians, researchers, and policy makers to develop relevant and accurate means for ensuring the health of this population.

Currently, evidence-based practices have gained popularity as a mechanism for reducing the burden of disease in the USA. While this is true for health and well-being in general within the USA, there is a dearth of research and evidence on assessments and treatment for ailments in the area of mental health. In fact, it has been well established that most of the "evidence" used to define evidence-based practices fails to adequately consider and address issues of racial, cultural, and socioeconomic diversity existent within the US population. So, though the creation of evidence-based interventions to improve mental health is an extremely important public health objective, too little of what we know about best practices in mental health (including assessment and treatment) applies to African Americans/Blacks specifically (Breland-Noble et al. 2011; Joe 2006; U.S. Department of Health and Human Services 2001). As indicated by the editors, accurate assessment of mental health concerns can facilitate improved pathways to treatment. In the case of mental health concerns, particularly for one of the leading causes of disability and time lost from work worldwide (depression), research demonstrates the utility of awareness and assessment as important facilitators of treatment for African Americans (Breland-Noble and Weller 2012).

Depression and suicide are of public health relevance given their established associations with reduced quality of life, reduced work and school productivity, implications for substance use and abuse, and loss of life via suicide. Though

---

A. M. Breland-Noble (✉)
Department of Psychiatry, Research Division, Georgetown University Medical Center, 2115 Wisconsin Avenue Suite 120, Washington DC 20007, USA
e-mail: ab2892@georgetown.edu

J. Sotomayor
Department of Psychology, Georgetown University, 306 N White-Gravenor Hall, 37th and O Streets, N.W., Washington DC 20057, USA
e-mail: jasonasotomayor@gmail.com

F. A. Burriss
Department of Social Work, North Carolina State University, 1911 Building, Campus Box 7639, Raleigh, NC 27695, USA
e-mail: faburris@ncsu.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_7, © Springer Science+Business Media New York 2015

depression *prevalence* is not an established health disparity between African American and white youth, research points to significant differences in depression prevalence and effective treatment access among African America and white adults (Alegria et al. 2008) as well as highly significant differences in depression treatment use and access between African American and white youth (Breland-Noble and AAKOMA Project Advisory Boards 2012; Breland-Noble et al. 2006, 2010). Pediatric and adolescent depression have emerged as public health concerns because of their associations with impaired functioning, poor decision-making, poorer quality of life, substance use, and diminished self-efficacy (Brook et al. 2010; Horwood et al. 2010; Okwumabua et al. 2003). With the enormity of the problem of depression in general which is exacerbated by health disparities in treatment for this disorder among African Americans, accurate assessment for this population is critical (Sashidharan et al. 2012).

Suicide is yet another unfortunate outcome often associated with depressive illnesses (Goldston et al. 2008). The rate of suicide within the African American community is the highest among adolescents and young adults with adolescent and young adult males exhibiting the highest rates of completed or attempted suicide. Crosby and Molock (2006) reported that for African Americans aged 15–19, suicide was the third leading cause of death, among those aged 20–29 years suicide is the fourth leading cause of death, and those aged 30–39, sucide is the eight leading cause of death.

During the early 1990s, suicide rates among African American males aged 15—24 years steadily rose and peaked in 1993 at 20.2, then began a steady decline to 11.6 (42.6 % decrease) in 2002 (Crosby and Molock 2006). In 2003, suicide was the 16th leading cause of death overall for African Americans and on an average day in the USA, one African American died by suicide every 4.5 h (Centers for Disease Control and Prevention (CDC) and National Center for Injury Prevention and Control 2005; Crosby and Molock 2006). Overall, there were 28,177 suicides recorded among African Americans from 1990 to 2003, and from 1999 to 2010 there were 16,466 suicides recorded among African Americans (CDC and National Center for Injury Prevention and Control 2005).

The estimated lifetime prevalence of suicidal ideation and attempts among blacks in the USA was 11.7 % and 4.1 % respectively, and among those with suicidal ideation (thoughts of death), 34.6 % made a suicide plan (Joe et al. 2006). Statistically significant differences ($P \leq 0.05$) were found between men and women, with attempts being more prevalent in women (4.9 % of the sample reporting), than men (3.1 % of the sample reporting). An ethnicity-by-sex analysis revealed the prevalence of suicide attempts was highest for Caribbean black men (7.5 %), followed by African American women (5.0 %). Caribbean black women had the lowest prevalence of attempts (2.7 %), while attempts were only slightly more prevalent for African American men (2.74 %) (Joe et al. 2006). Joe and colleagues also found that the 12-month prevalence rate for suicidal ideation (12.8 %) and nonfatal suicide attempts (5.0 %) in African American women is high relative to men and women of other ethnic groups.

Overall, it is well established that African Americans are susceptible to mental illnesses like depression and events like suicide (even though suicide is a rare occurrence within this overall population). Given established prevalence rates (which are statistically comparable among youth yet statistically different among adults), under treatment of the disease (and event), and the associated impairment and loss of life, diagnosis is key as a first step toward intervention. Though a multitude of tools are available for use in assessing depression and suicidality in youth and adults, very few have empirical data associated with them to support their use with African Americans and people of African descent. Further, we identify the limited number of tools evaluated and/or used with African American populations to assess depression and suicidality and describe their prior use with African Americans. We begin with adult measures followed by child and adolescent measures.

Adult Depression Measures

The Beck Depression Inventory (BDI) The BDI is a 21-item multiple-choice inventory designed to allow patient report of the types and severity of experienced depressive symptoms for the week prior to the date of the assessment (BDI-original) and the prior 2 weeks from the date of assessment (BDI-II) (Beck et al. 1996; Brown et al. 2000). The BDI focuses on the cognitive aspects of depression symptomology (i.e., what a person's thoughts and perceptions are regarding their depression). The BDI-II is a 21-item revision of the BDI, which focuses on the symptoms of depression as they are described in the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV). This revision was based on substantial revisions to the clinical conceptualization of depression included in the DSM-IV (Beck et al. 1996; Brown et al. 2000). Though the BDI and BDI-II are technically not the measures of suicidality, they both do contain a single item assessing suicidal ideation. While the depression items of the BDI and BDI–II focus on depression, they address different aspects of depression; the BDI addresses cognitive aspects while BDI-II addresses symptomatology. Specific details regarding research on the standardization and psychometric properties of the BDI are provided below. As the reader will see, the BDI (both versions) is a widely used and accepted inventory for the assessment of depression. Based on its widespread use and validation and psychometric property assessment in African American and Afro-Caribbeans, it seems a strong measure for use with African Americans and Blacks of the African diaspora.

Standardization Samples The BDI and BDI-II have been collectively examined using extensive studies to assess their utility, accuracy, and psychometric properties with people across the lifespan (Aaron et al. 1988). This depth of research on the normative properties is rare among clinical assessment measures but does allow one to create a picture of how well the BDI may function with African American and African diasporic populations. Sashidharan (2012) examined the BDI in a sample of university undergraduates that included 139 African Americans matched to an equal number of white students who were randomly selected from a larger sample of over 800 white students. The research team reported no statistically significant differences between African American and White students' mean scores on the BDI-II. Further, they reported no statistically significant correlations between BDI-II scores and race. Additional studies of the psychometric properties of the BDI include significantly under-resourced and socioeconomically disadvantaged African Americans as well as the middle class, college educated Afro-Caribbean populations but consistently report strong psychometric properties (as elucidated in the next section; Campbell et al. 2012; Grothe et al. 2005; Joe et al. 2008; Kneipp et al. 2009)

Psychometric Properties Both versions of the BDI appear to be beneficial for helping clinicians assess symptoms of depression as well as their severity in African American populations. As an example, we note the work of Grothe whose analyses (in a sample of medical outpatient African Americans) indicated that individuals with a diagnosis of current major depression had significantly greater BDI-II total scores ($M = 23.12$, $SD = 8.66$) than those without a depression diagnosis ($M = 8.23$, $SD = 7.50$); t (218) = 12.83, $p < 0.01$ (Grothe et al. 2005).

Criterion validity has been demonstrated in the BDI-II via multiple studies differentiating depressed from nondepressed individuals. Specifically, Arnau and colleagues (2001) demonstrated the ability of the BDI-II to differentiate between depressed and nondepressed primary care patients. Additionally, using the PRIME-MD, Grothe and colleagues (Grothe et al. 2005) established the criterion validity of the BFI-II in their outpatient, low-income African American sample. Suicide items on the BDI has also been examined for criterion validity with findings establishing its strong properties. In fact, the BDI suicide item was moderately correlated (rs = 0.56–0.58) with the Beck Scale for Suicide Ideation for a psychiatric sample (Beck and Steer 1991).

Discriminant validity of the BDI was demonstrated through weaker correlations (0.71, $p < 0.001$) with measures of anxiety than with other measures of depression (0.89, $p < 0.001$; (Steer et al. 1997). Joe and colleagues found strong evidence for the convergent validity of the BDI-II (Joe et al. 2008) and the Hamilton Depression Rating Scale (the HAM-D) in a sample of under-resourced African Americans (Joe et al. 2008). The predictive validity of the BDI-I (in relation to both suicide and depression) has been demonstrated in multiple studies using both community sample of adolescents and adults. In each case, the BDI suicidal ideation item was found to predict future suicide attempts (OR = 6.9) and future depressive episodes (OR = 2.1; (Lewinsohn et al. 1994). Additionally, using data from the Brown et al.'s (2000) study, it was found that individuals who scored a 2 or higher on the BDI suicide item were 6.9 times (95 % CI: 3.7–12.6) more likely to commit suicide than those who scored less than 2 (Brown et al. 2000).

Using samples of African Americans, the internal consistency of the BDI–II and its factors is indicated to be quite high with a BDI–II total score, $\alpha = 0.90$; cognitive factor, $\alpha = .81$; and somatic factor, $\alpha = 0.87$ (Grothe et al. 2005). Separate studies showed the BDI as possessing a Cronbach's alpha of $\alpha = 0.89$ and an internal consistency reliability of $\alpha = 0.94$, a test-retest reliability of 0.75 and a split-half reliability coefficient of 0.90 (Joe et al. 2008). Consistent with Grothe, Sashidharan (2012) found evidence to support the dimensionality, internal reliability, and convergent validity of the BDI–II in a sample of African American participants.

The Center for Epidemiology Studies Depression (CES-D) Scale The CES-D Scale is a 20-item self-report measure developed by the National Institute of Mental Health for the assessment of depressive symptom prevalence in the general population (Radloff 1977). Though it is not a diagnostic tool for mental health clinicians, it is quite useful for helping clinicians understand the type and frequency of *affective* depressive symptoms that a patient might experience. Questions are answered on a 4-point Likert scale where responses range from "rarely or none" to "most or all" of the time. CES-D scores range from 0 to 60, with a score of 16 indicating the presence of clinically significant depressive symptoms (Radloff 1977).

While limited research exists (which is detailed below) examining the CES-D exclusively with African Americans, the available data does point to the utility of the measure in helping researchers and clinicians accurately assess the presence of depressive symptomatology in this population. Though the CES-D is not recognized as a diagnostic screening tool for clinical use, it can be useful in helping clinicians gain insight into the clinical presentation of symptoms experienced by depressed African Americans.

Standardization Sample and Psychometric Properties Fortunately, the CES-D standardization sample included approximately 25 % ($n = 259$) African Americans, though this sample was significantly poorer than the white sample with a mean annual household income approximately 40 % lower than that of their white counterparts (Comstock and Helsing 1977).

Since the initial validation of the CES-D, the psychometric properties of the measure have been examined with African American samples of differing socioeconomic strata. Rozario and colleagues found moderate to strong internal consistency in a sample of low-income African Americans (0.83) while Williams and colleagues (Williams et al. 2007) found an even lower yet acceptable level of internal consistency in a sample of well-educated middle class African American women (0.75) (Rozario and Menon (2010). Most of the current data on the utility and psychometric properties of the CES-D appears to come from samples of under-resourced (i.e., low-income) African Americans (Nguyen et al. 2004).

Radloff (1977) identified an internal consistency reliability of approximately 0.85 in the general population and about 0.90 in a patient sample using the CES-D. She also found test-retest reliabilities of: 0.51 after 2 weeks, 0.67 after 4 weeks, 0.59 after 6 weeks, 0.59 after 8 weeks, 0.48 after 3 months, 0.64 after 6 months, and 0.49 after 12 months in a sample of participants.

The CES-D scale correlates strongly with other self-reported depression measures as well as variables related closely to the clinical diagnosis of depression. In a racially diverse sample (primarily African American and Latino) of over 200 Traumatic Brain Injury patients, the CES-D was found to have moderate to strong concurrent validity when compared with the Beck Depression Inventory ($r = 0.67$) (Bush et al. 2004). Researchers have found that even within African American subjects, CES-D scores are much higher for depressed vs. nondepressed patients (Nguyen et al. 2004).

Primary Care Evaluation of Mental Disorders (PRIME-MD) The PRIME-MD was the first instrument designed for use in primary care settings to screen for specific mental disorders using criteria from the DSM-III-R and DSM-IV (Spitzer et al. and The Patient Health Questionnaire Primary Care Study Group 1999; Spitzer et al. 1994). The PRIME-MD consists of the 26-item patient self-administered questionnaire (PRIME-MD-PQ) and a clinician administration evaluation guide PRIME-MD-CEG), to guide clinicians in more detailed data collection via five modules designed to more fully elucidate the data generated by the PRIME-MD-PQ (Tamburrino et al. 2009). The PHQ-9 is the depression specific subscale of the PRIME-MD and has been examined extensively for its utility in screening for depression symptomatology and severity. Overall, the PRIME-MD is potentially useful for persons working with racially diverse patient populations as research points to the primary care setting as the main point of entry for African American adults with mental health concerns (Snowden and Pingatore 2002; Spitzer et al. 1999). Though, we briefly describe the PRIME-MD to provide context, our focus is on the PHQ-9 as it is most relevant for depression screening.

Standardization Sample and Psychometric Properties The PRIME-MD was validated in racially diverse sample of 1000 patients via 31 physicians across four primary care clinics in the USA (Spitzer et al. 1994). The study found the measure to be an accurate screener for depression and four other major mental illnesses typically seen in primary care settings. The measure is quickly administered (i.e., it generally takes less than 10 min to complete) and based on the original study, diagnoses were accurately confirmed by an independent panel of mental health professionals with an overall accuracy rate of 88 %. The PHQ Depression screen was part of the original PRIME-MD and remains so today. Currently entitled the PHQ-9, this depression screen consists of nine questions scored on a Likert scale.

PRIME-MD has not been specifically normed for African Americans, particularly African American women; however, the measure does appear to be accurate with diverse populations (Miranda et al. 2005)—13 % of the participants in the PHQ-9 Primary Care study were African American (Ramos et al. 2004; Spitzer et al. 1999). With 92 % accuracy for reporting major depression (Miranda et al. 2005; Ramos et al. 2004; Spitzer et al. 1994), PRIME-MD reports good agreement with independent psychiatric diagnoses guided by a structured interview.

Patient Health Questionnaire-9 (PHQ-9) The PHQ-9 is a self-administered version of the PRIME-MD comprised of the PRIME-MD-PQ and PRIME-MD-CEG, the patient questionnaire and the clinician evaluation guide (Kroenke and Spitzer 2002). Overall, the full PHQ allows patients to respond to questions reflecting eight DSM-IV specific diagnoses including panic disorder, anxiety disorders, eating disorders, and major depressive disorder (Kroenke et al. 2001; Spitzer et al. 1999). The depressive disorder subscale of the overall PHQ is entitled the PHQ-9 and is a measure that supports the diagnosis of depression along with depression symptom severity. Scoring for individual items is on a Likert scale of “0” (not at all) to “3” (nearly every day), yielding an overall score range of 0–27. For diagnoses of major depression, a respondent must endorse five or more of the nine depressive symptoms for at least “more than half the days” in the prior 2 weeks with a required symptom of depressed mood (Kroenke et al. 2001). Given that this measure has been used with large sample of under-resourced African American women (see below for details), it may be useful for clinicians and researchers working with African American populations. Future studies would however need to replicate findings in more socioeconomically diverse samples.

Psychometric Properties Data from both the PHQ Primary Care Study (Spitzer et al. 1999) and the PHQ Obstetrics-Gynecology Study (Kroenke et al. 2001; Spitzer et al. 2000) provide the following strong evidence of measure validity. First, using a sample of 580 primary care patients who completed the PHQ-9 and who were subsequently reinterviewed by mental health professionals, researchers were able to demonstrate the criterion validity of the measure. Next the strong association between PHQ-9 scores and functional status, disability days, and symptom related difficulty was seen in two validation studies of the PHQ-9 conducted in multiple primary care and Obstetrics-Gynecology clinics across the USA. These validation studies, including the Obstetrics-Gynecology study with a subsample of approximately 450 African American women helped to establish the construct validity of the PHQ-9 and yielded Cronbach's α of 0.89 and 0.86 respectively (Spitzer et al. 1999; Spitzer et al. 2000).

Summary

Overall there are a limited number of depression assessment measures available that have been specifically developed and/or utilized extensively with African American adults. Of the available measures, we have described those (which in our view) have the most applicability and are the most easily accessible to clinicians, researchers, and policy makers interested in working with African Americans.

Suicide Assessment

We now turn our attention to the assessment of suicidality in African American adults by examining a number of measures, as well as subscales of depression measures.

The Scale of Suicide Ideation (SSI) The SSI is a 21-item, interviewer-administered rating scale that measures the current intensity of individual's specific attitudes, behaviors, and plans to commit suicide on the day of the assessment (Beck et al. 1979). Each question is answered on a 3-point scale, from 0 to 2. A total score for the SSI is based on the first 19 items, for a score range of 0–38. The SSI includes five screening items assessing an individual's level of desire toward a suicide attempt and the wish to live or die. If a patient reports a passive or active suicide wish, an interview may complete 14 additional questions with the patient to examine specific suicidal risk factors, including the duration and frequency of ideation, an individual's feelings of self-control regarding making an attempt, the number of deterrents available to the patient, the incidence and frequency of prior attempts, and the activity engaged in toward a contemplated attempt (Brown and Goldston 2000).

Additionally, the normative sample for the SSI included psychiatric inpatients and outpatients with a relatively sizeable proportion of African Americans overall (Beck et al. 1985, 1997). The measure has strong psychometric properties including in samples of African Americans with Cronbach's alphas for racially diverse samples ranging from 0.84 (Beck et al. 1997) to 0.89 (in a predominantly African American sample) (Harris and Molock 2000). The measure is reported to demonstrate concurrent validity with the suicide item from the Hamilton Rating Scale for Depression, reports of previous suicide attempts and depression severity (Beck et al. 1979, 1985, 1997). Given these study outcomes, the SSI appears to be a solid tool for clinician use with African American patients.

The Suicide Probability Scale The Suicide Probability Scale is a 36-item measure of current suicidal ideation, negative self-concept, hopelessness, and negativity/hostility (Cull and Gill 1988). Individual items are scored on a Likert scale of 1 ("none or little of the time") to 4 ("most or all of the time") points. The measure generates three summary scores; a suicide probability score, a total weighted score, a T-score and four subscale scores including hopelessness, suicidal ideation, negative self-evaluation, and hostility.

The SPS was standardized on a multiracial sample of inpatient and outpatient adults and adolescents including a sample that was almost 15 % African American (Brown 2002; Cull and Gill 1988). The properties of this measure are strong with Cronbach's alpha reported at 0.93 and test-retest reliability of 0.92 over a 3-week period (Cull and Gill 1988). Additionally, the measure demonstrates good concurrent validity with the suicide threat scale of the MMPI—Minnesota Multiphasic Personality Inventory (Cull and Gill 1988; Goldston et al. 2008).

While the SPS has adequate psychometric properties, it has not been evaluated for its consistency with the standard definitions of suicidality as defined in a seminal article by O'Carroll and colleagues (Ocarroll et al. 1996). Overall, the SPS may require additional testing in samples of African Americans as well as for its concurrent validity with standard measures of suicidality before it is used with African Americans. Specifically, since research indicates trends toward African American and white differences in the perception, expression, and severity of hopelessness and hostility (Maier et al. 2009), this measure could benefit from additional examination of these specific constructs in African American samples (particularly boys and men).

Suicide Subscales of Depression Measures

Beck Depression Inventory The Beck Depression Inventory (both versions) include a set of subscales focused on the assessment of suicidality. The primary difference between the subscales of the two versions is the time frame, with the BDI focusing on the prior week and the BDI-II focusing on the prior 2 weeks from the day of the assessment. On either inventory, individuals are asked to decide which of the following best describes the way they have been feeling: (1) "I do not have any thoughts of killing myself," (2) "I have thoughts of killing myself, but I would not carry them out," (3) "I would like to kill myself," and (4) "I would kill myself if I had the chance." Any individual who rates this final item 4 with a 2 or higher indicates his/her intent to die (Brown 2002). Given prior examinations of the BDI overall in African American/Black populations, it seems reasonable for clinicians to use this measure and feel comfortable that it provides accurate assessments of suicidality. However, we caution clinicians to consider additional means of qualitative inquiry to supplement the use of the measure and provide a fuller picture of the nature and extent of suicidality among African Americans/Blacks.

Child Measures

The Children's Depression Inventory (CDI) The CDI is a self-report measure for use with children and adolescents aged 7–17 (Kovacs 1985). Dr. Kovacs created the measure to address concerns over the use of the Beck Depression Inventory (BDI) with younger populations. The measure includes 27 sets of items comprised of three sentences that describe a young person's thoughts or feelings over the 2 weeks prior to taking the inventory. Respondents are asked to choose which single sentence describes them best (Goldston 2000; Kovacs 1985). The CDI yields a total score ranging from 0 to 54, as well as five subscores: Mood, interpersonal problems, ineffectiveness, anhedonia, and negative self-esteem. Scores of 19 or higher indicate clinical significant depression. Although the CDI does not provide a direct assessment of suicidality, it does include a single item assessing suicidal ideation.

The Children's Depression Inventory II is a full revision of the original Children's Depression Inventory completed in 2010. The new features of the measure include, "new items that focus on the core aspects of childhood depression, revised scales, and new norms that are representative of the US population." (Kovacs 2010; Multi-Health Systems Inc. 2011, p. 1). The measure was designed as a comprehensive measure to assess depressive symptoms in youth aged 7–17 and is available in three forms. The three forms include a child self-report version (CDI-2 Self-report), a parent version (CDI-2 Parent), and a teacher version (CDI-2 Teacher). This version yields a total score, 2 scale scores (focused on emotional and functional problems) and 4 subscale scores (focused on negative mood/physical symptoms; negative self-esteem, interpersonal problems; and ineffectiveness) from a series of questions to which respondents answer " 0 (absence of symptoms), 1 (mild or probable symptom), or 2 (definite symptom)". The parent version includes just 17 items of "observable" depressive behavior and symptoms to which parents/legal guardians respond using a 4-point scale (0 (not at all), 1 (some of the time), 2 (often), or 3 (most of the time)). The CDI-2 Parent form yields a total score and the 2 scale scores.

Standardization Sample and Psychometric Properties Only two of the approximately 5–6 studies focused on establishing the psychometric properties of the CDI include significant numbers of African American youth. In the first study, Carey and colleagues included a sample that was approximately 38 % African American (Carey et al. 1987) while in another, the sample of African American youth was approximately 38 % of psychiatric inpatient youth (Hodges 1990).

McLaughlin, Hilt, and Nolen-Hoeksema (2007) tested the CDI's reliability with a sample of middle-school aged children comprised of 11.8 % non-Latino African Americans, while another study consisted of children aged 7–12, of which 23.1 % were children of color (Finch et al. 1987). A separate study divided groups in emotionally disturbed and nonemotionally disturbed children. About 25 to 40 % of the children in the emotionally disturbed group were children of color, while approximately 33.3 % of the nonemotionally disturbed group were children of color (Saylor et al. 1984). A separate study included a clinical sample and a nonreferred sample that were 38 and 39 % African American, respectively (Carey et al. 1987). The last study reviewed in this section sampled inpatients between the ages of 6 and 18, of which 28.7 % were African American (Nelson et al. 1987).

Administering the CDI among adolescents representing different racial/ethnic groups to examine the differences in both internalizing and externalizing symptoms, McLaughlin et al. (2007) found there to be good internal consistency reliability in the CDI ($\alpha = 0.82$), as well as among each racial/ethnic group represented: white ($\alpha = 0.82$), black ($\alpha = 0.79$), Hispanic ($\alpha = 0.82$), and other ($\alpha = 0.85$). Nelson et al. (1987) found a similar Cronbach's alpha ($\alpha = 0.86$) and also calculated that the CDI's inter-item reliability coefficient ranged from 0.021 to 0.435. Test-retest reliability was calculated at 0.87 ($p < 0.001$) after 1 week, and at 0.59 ($p < 0.006$) after 6 weeks (Saylor et al. 1984), while a different study found the CDI's test-retest reliability to be 0.82 ($p < 0.001$) for a 2-week interval and between 0.66 and 0.67 ($p < 0.001$) for longer intervals (Finch et al. 1987). Okwumabua et al. (2003) found that the distributional characteristics of the CDI reported in their sample were similar to those in previous studies with low-income African American youth.

Carey and colleagues (1987) examined the construct and discriminant validity of using the principal component and discriminant analyses, in a large sample of inpatient psychiatric/residential subjects and nonreferred subjects. They determined that the factor structure of the CDI remained generally stable across the populations examined. Moreover, they concluded that clinical and nonreferred subjects could be reliably distinguished using CDI factor scores.

Concurrent validity was examined by predicting whether or not the CDI could reliably identify children meeting DSM diagnostic criteria for depression on the Child Assessment Schedule (CAS) via sensitivity and specificity calculations. Children meeting criteria for depression per the CAS were compared to those labeled as depressed according to the CDI. The CDI had a specificity of 84 % (i.e., the incidence of true negative results attained from the CDI when depression is absent). Unfortunately the sensitivity of the CDI was a mere 54 % representing the proportion of cases in which the CDI made a positive diagnosis when depression was present (Hodges 1990).

These findings provide evidence for the convergent and discriminant validity of the CDI in that depressed children score significantly higher on the CDI than nondepressed children. In contrast, neither the conduct-disordered nor the anxiety-disordered children scored significantly higher on the CDI than children without these specific diagnoses, which lends support to the discriminant validity of the CDI (Hodges 1990).

Recommendations The CDI is a widely used child depression measure with some demonstrated adequate psychometric properties for African American youth and most cross-cultural research on the initial form of the measure with international white or unspecified populations (Allgaier et al. 2012). Overall it seems better suited as a screening measure and may not be suitable as a diagnostic tool. Of the research available, very little points to the utility of the revised form of the measure, published in 2010, with African American youth across the socioeconomic spectrum. Future research should therefore examine the use of this measure on socioeconomically and racially diverse samples of youth.

The Children's Depression Rating Scale (CDRS)

The Children's Depression Scale (CDRS) is a 16-item, clinician-administered measure (later revised to a 17-item measure—CDRS-R; Poznanski and Mokros 1996) for the assessment of depression in children between the ages of 6 and 12 years old. Items are measured on a 3-, 4-, 5-, and 6-point scale, and inquire about social withdrawal, capacity for fun, sleep habits, eating habits, irritability, schoolwork, expressive communication, general somatic features, hypoactivity, and depressed mood (Poznanski et al. 1979). The revised version has a score of ≥ 40 as indicating depression, whereas a score of ≥ 28 represents minimal or no depressive symptoms (Poznanski et al. 1979).

Standardization Sample and Psychometric Properties The CDRS was devised and tested on 30 inpatient children in a medical hospital (Poznanski et al. 1979). A follow up to this study was conducted to examine the preliminary psychometric properties of the CRDS-revised version in a sample of 53 outpatient youth 43 % of whom were African American (Poznanski et al. 1984). The utility of the revised measure was also examined in a group of 233 school aged children of unspecified race (Poznanski et al. 1985). Finally, in a separate study examining the psychometric properties of the CDRS-R, adolescents with symptoms of depression were assessed. Of all the participants, only 10.3 % were African American (Mayes et al. 2010).

Research is equivocal regarding the utility of the CDRS with African American youth specifically. Adolescents with symptoms of depression were assessed with the CDRS-R and the Children's Global Impressions Severity (CGI-S) at screening, baseline, and after 12 weeks of fluoxetine treatment. Internal consistency for the CDRS-R was good at all three stages (screening: $\alpha = 0.79$; baseline: $\alpha = 0.74$; exit: $\alpha = 0.92$). Construct Validity was found in the total score being highly correlated with global depression severity ($r = 0.87$, 0.80, and 0.93, respectively; $p < 0.01$) (Mayes et al. 2010). However, Stein and colleagues found that the clinician administered CDRS functions differently across racial groups in that clinicians rated African American (and Latino) youth as demonstrating more severe symptoms on the observational items of the CDRS-R compared to their white peers. This type of findings may warrant future study to determine if any attributional bias exists in clinician ratings of African American youth depressive symptoms.

Recommendations As symptoms of depression often overlap with other psychiatric disorders, such as other mood, anxiety, and personality disorders, it would be wise to investigate ways in which the collection of psychometric data can be done while holding constant the potential presence of symptoms from other psychiatric disorders. Additionally, since the CDRS-R was only tested with small samples of African American youth, further research must be done to ensure that the CDRS-R is sufficiently effective in measuring depression specifically in African Americans.

The Reynolds Adolescent Depression Scale (RADS) The RADS is a 30-item self-report measure for adolescents between the ages of 13 and 18. The RADS measures depression on five different levels (somatic, anhedonia, cognitive, negative view of self, and loneliness) and does so on a four-point (1–4) Likert scale, where 1 = almost never, 2 = hardly ever, 3 = sometimes, and 4 = most of the time. The cutoff score has been established at 77, where any score above 77 indicates a symptom endorsement associated with clinical depression. Though, there is one suicide item on the RADS ("I feel like hurting myself"), the RADS is not meant for assessing suicidal ideation or behaviors, but rather to assess the symptoms associated with depression (Reynolds 1986). Though, a second version of the RADS, the RADS-2 (Reynolds 2002), has been developed, it is largely similar to the original RADS, so we will focus mostly on the original version.

Standardization Sample and Psychometric Properties The RADS was standardized on adolescents from 7th to 12th grade (Reynolds 1986). Molock, Puri, Matlin, and Barksdale (2006) conducted a study using the RADS with 212 African American high school students, whose ages ranged from 13 to 19, from a suburb of Washington DC. In another study, subjects were adolescent inpatients from 12 to 17 years old, of which 12 % were African American (Krefetz et al. 2002). A final examination of the RADS included adolescent psychiatric inpatients (14–17 years old), of which 12.2 % were African American (Osman et al. 2010).

The test developer has consistently found high internal consistency reliability ($\alpha s > 0.90$) and high test-retest reliability ($rs > 0.60$) (Reynolds 1986). Molock et al. (2006) also found a Cronbach's alpha of 0.90. The inter-item reliability coefficient ranged from 0.909 to 0.939, while the split-half reliability coefficient was 0.91 (Reynolds 1986). Conventional Cronbach's alphas for the RADS-2 were calculated for the Dysphoric Mood, Anhedonia-Negative Affect, Negative Self-Evaluation, Somatic Complaints subscales, and the total RADS-2 as 0.85, 0.80, 0.87, 0.80, and 0.94, respectively (Osman et al. 2010).

Concurrent validity was shown between the RADS and the Hamilton Depression Rating Scale with a correlation coefficient of $r = 0.83$ (Reynolds 1986). Convergent validity was demonstrated by strong correlation coefficients with other self-report measures of depression, such as the BDI and the CDI (Reynolds 1986).The strong positive correlation ($r = 0.84$, $p < 0.001$) between the BDI-II and the RADS is another indicator of strong convergent validity (Krefetz et al. 2002). Construct validity was provided by several studies, in which the RADS was compared to other self-report measures for depression, such as the BDI, the CES-D, the Self Rating Depression Scale (SRDS), and the CDI. Correlations between the RADS and each of these other measures were 0.73, 0.75, 0.72, and 0.73, respectively (all of which had $p < 0.001$).

Recommendations Other than Molock et al. (2006), the other studies reviewed here for the RADS have minimal African American participation since the samples used were predominantly white so the findings illustrated in this review should be taken with caution.

Suicide Measures for Children and Adolescents

The Suicidal Ideation Questionnaire—Junior (SIQ-JR) The Suicidal Ideation Questionnaire (SIQ) is a screening measure for the degree of "seriousness" of suicidal ideation. Two self-report versions of the SIQ have been developed: a 30-item version (called the SIQ) originally designed for adolescents in the 10th, 11th, and 12th grades, and a 15-item version (named the SIQ-JR) originally designed for adolescents in grades 7, 8, and 9. Respondents to both versions of the SIQ answer each of the items on a 7-point scale, ranging from 0 ("I never had this thought") to 6 ("almost every day."). Scores range from 0 to 90, with a published clinical cutoff score of 31 (Goldston 2000; Reynolds 1987).

The questions in both versions of the SIQ are founded on theoretical notions of a hierarchy of severity of suicidal behavior and cognitions. This hierarchy states that suicidal behavior and thoughts form a continuum ranging from just thoughts of death, to thoughts of wanting to be dead, to general and then specific thoughts of suicide, to making specific plans for suicidal behavior, all the way to actually attempting to kill oneself. As most suicide-related items, the wording of suicide ideation questions implicitly refer to "nonzero intent to die" (O'Carroll et al. 1996; Reynolds 1987). There is no item on past or current suicide attempts, so the SIQ cannot be used to identify suicide attempters.

The SIQ-JR has been used in several studies with African American representation in the samples, and has shown great reliability and validity. Though, Reynolds (1987) has insisted that it not be used as a predictive measure for suicide, the SIQ has proven to be useful for identifying suicidal ideators and suicide attempters. Relevant research is summarized below.

Standardization Sample and Psychometric Properties One study had an adolescent inpatient sample consisting of 87 % white youth (King et al. 1995), while another study had 13 % of their adolescent (ages 13–18) inpatient sample as African American (Spirito et al. 1987). Pinto, Whisman, and McCoy (1997) conducted a representative sample of socioeconomic status but had no normative data of the cultural groups from which they sampled. Reynolds and Mazza (1999) examined the SIQ in adolescent participants from middle schools, of which 91.2 % were African American or Hispanic.

The internal consistency reliability of the SIQ-JR within the standardization sample was $\alpha=0.91$ with similar reliability coefficients found for both males and females, $\alpha=0.90$. On the second assessment of the SIQ-JR, the internal consistency reliability was $\alpha=0.94$ for the total sample with again similarly high reliability coefficients found for males ($\alpha=0.92$) and females ($\alpha=0.93$) (Reynolds and Mazza 1999). Reynolds (1987) calculated a high internal consistency reliability, $\alpha=0.97$. Pinto et al. (1997) also found a high Cronbach's alpha for the SIQ-JR ($\alpha=0.97$). Results of the test-retest reliability analyses of the SIQ-JR for the total sample and by gender show a coefficient of 0.89 for the entire sample, a coefficient of 0.87 for females, and a coefficient of 0.93 for males—all of which are very high (Reynolds and Mazza 1999). Reynolds (1987) also found high test-retest reliability in his standardization sample of adolescents, at 0.72 over 4 weeks.

Pinto et al. (1997) found that adolescents who had attempted or ideated about suicide did not differ with respect to scores on the SIQ, yet both groups had higher scores that nonsuicidal adolescent inpatients. Relatedly, Spirito et al. (1987) observed that, within their adolescent pediatric inpatient sample, suicide attempters who were assessed as having chronic psychiatric problems had higher SIQ scores than those with acute psychiatric problems. Both findings are indicative of the SIQ's concurrent validity. Regarding the measure's predictive validity, though the SIQ is not meant for the prediction of suicide, King et al. (1995) found that SIQ-JR scores from their adolescent inpatient sample were predictive of later suicide attempts.

The validity of the SIQ-JR was determined through a contrasted group approach. This was done through the examination of the SIQ-JR scores of adolescents who had attempted suicide compared to those who had not, given that a history of attempted suicide has been found to be a significant risk factor for later suicidal behaviors. Adolescents who had attempted suicide reported significantly higher levels of suicidal ideation on the SIQ-JR ($M=28.76$, $SD=22.34$) than those who had not attempted suicide ($M=7.91$, $SD=8.37$) (Reynolds and Mazza 1999).

A recency effect was also anticipated, in the sense that the recency of suicide attempt would be correlated to higher SIQ-JR scores. Adolescents who admitted a suicide attempt in the previous 12 months reported a significantly higher SIQ-JR score ($M=43.00$, $SD=22.94$) than did adolescents who attempted suicide more than a year previously ($M=18.80$, $SD=16.40$) (Reynolds and Mazza 1999). An analysis of variance was done to further investigate the relationship between SIQ-JR scores and suicide attempt history between the groups of adolescents who reported no history of suicide attempts, an attempt more than a year previously, and an attempt within the past 12 months, which resulted in significant differences between all three groups.

The Diagnostic Interview Schedule for Children (DISC) The DISC is a structured psychiatric diagnostic interview for youth and their parents, covering a myriad of mental disorders seen in children and adolescents ages 6 through 18. While this measure is not depression-specific, it was created to mirror similar adult assessment tools. The latest revision of the DISC by the National Institute of Mental Health (NIMH DISC-IV) is based on the diagnostic criteria from the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV). However, an earlier version of the DISC, the NIMH DISC–2.3, based on the DSM-III-R, is also still in use due to the fact that the most salient differences between the two is that the DISC-IV modified certain unreliable questions and added new sections for schizophrenia and substance abuse—none of which affected questions surrounding mood disorders (Shaffer et al. 2000). Thus, both DISC versions will be reviewed throughout this section.

The DISC is also available in a computer-assisted version, the C-DISC. Lay interviewers can administer this measure after 1–2 days (for the computerized version; 4–5 days for the written version) of training since most measure questions are administered verbatim and worded so that respondents answer via a simple "yes," "no," "somewhat," or "sometimes" (Goldston 2000; Shaffer et al. 2000).

Depressive symptom queries of the NIMH DISC-IV are similar, and reference the 2 weeks, 4 weeks, and a year preceding the interview (Goldston 2000; Shaffer et al. 2000). Like its predecessor, the NIMH DISC-IV focuses on depressive symptoms, but includes questions regarding lifetime suicide attempts, as well as those in the 4 weeks and in the last year immediately prior to the interview. The question about age of first suicide attempt (in the NIMH DISC 2.3) is not included in the NIMH DISC-IV, replaced instead with a question about whether reported suicide attempts required medical attention.

Neither version of the DISC assesses non-suicidal self-injurious behaviors, like cutting. Goldston (2000) argues that the query regarding suicidal ideation in the Youth and Parent versions of the DISC are likely to generate conservative estimates of suicidal ideation due to the use of the word "seriously" as a context for the ideation question (Goldston 2000; Shaffer et al. 2000). Since this word "seriously" is not well explicitly defined, Goldston (2000) argues that the word conveys ambiguity for interviewees, making it difficult for them to respond accurately.

Standardization Sample and Psychometric Properties The community sample used to establish norms for the DISC are described in (Shaffer et al. 2000) as composed of 60% African Americans/Hispanics, while the (Gould et al. 1998) community sample was 13.8% African American. To examine the test-retest reliability of the DISC, one study collected a community sample of parents and children recruited from child and adolescent psychiatric outpatient clinics. All subjects were clinically diagnosed over the past year with DSM criteria for certain common disorders, including depressive disorders. The NIMH DISC-IV was administered twice by lay interviewers. A separate study conducted a similar reliability test using the DISC–2.1. The DISC-IV is regarded as more reliable than previous versions of the DISC with respect to depressive disorders (Shaffer et al. 2000).

Recommendations Though the DISC is one of the more widely recognized assessments used to assess depression in children and adolescents in research studies in particular, relatively few studies have examined its validity and reliability for African Americans. Therefore, further tests on the psychometric of the DISC are encouraged before allowing its extensive use within this cultural groups.

Assessment of Youth Suicide (Subscales)

The Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADS) Though there are several versions of the K-SADS, our focus here is on the version that has been used most extensively in research over the years—the K-SADS, School Age-Epidemiologic Version (K-SADS-E). The K-SADS-E is a semi-structured interview performed by clinicians on children and adolescent between the ages of 6 and 18, and their parents. Among other psychiatric disorders, the K-SADS-E includes a major depression section in which items regarding suicide and suicidal ideation can be found and have been used in depression and suicide research (Goldston 2000; Orvaschel and Puig-Antich 1987).

What differentiates the K-SADS-E from the other versions of the K-SADS is the specificity in questions about suicidal behaviors. The K-SADS-E has items revolving around recurrent thoughts of wanting to die, suicidal ideation (“Have you thought about hurting or killing yourself?”), the presence of a plan for suicide, suicide attempts (“Did you ever try to kill yourself or done anything that could have killed you?”), and even nonsuicidal physically self-damaging suicidal behaviors (Goldston 2000; Orvaschel and Puig-Antich 1987). The question dealing with suicide attempts is straightforward and implies the standard nonzero intent to die (O’Carroll et al. 1996).

Standardization Sample and Psychometric Properties Two studies used the same group of adolescents aged 13–18 in high schools. Of these adolescents, the overwhelming majority was Caucasian (91.1 %), with no indication of the races of the other 8.9 % (Lewinsohn et al. 1993; Lewinsohn et al. 1994). Rohde, Mace, and Seeley (1997), on working with adolescents, only had about 1.7 % of their participants as African American. A separate study found the vast majority of their participants to be adolescent males (92 %) and Caucasian (92 %) (Renaud et al. 1999). The last study we examined for standardization properties for the K-SADS-E gave no demographic information on the race of their participants (Frazier et al. 2007).

Lewinsohn et al. (1993) used the K-SADS-E to see how adolescents differed if they had or had not attempted suicide in the past. Overall, they found that adolescents with a history of suicide attempts had: more use of medications, worse health, worse coping skills, lower self-esteem, greater current and lifetime suicidal ideation, more pessimism, more negative attributions, and a greater chance of being diagnosed with depression. In a separate study, it was found that current and lifetime suicide attempts assessed with the K-SADS-E were correlated with a higher chance of being diagnosed with major depression, dysthymia, and anxiety disorders (Rohde et al. 1997). Additionally, in a retrospective psychological autopsy study in which the K-SADS-E was used, suicide completers were discovered to have had a more extensive history of nonlethal suicide attempts than matched community controls (Renaud et al. 1999).

With regards to the predictive validity of the K-SADS-E, Lewinsohn et al. (1994) noted that prior history of suicide attempts is one of the strongest predictors of subsequent suicide attempts, astonishingly increasing the risk 18-fold. Lastly, Frazier et al. (2007) established a high inter-rater reliability for the K-SADS-E for mood disorders across age groups (ages 4–10: $k=1.00$; ages 11–17: $k=0.95$).

Recommendations As both Lewinsohn et al. (1993, 1994) studies relied heavily on statistical comparisons, type I error was rather high, and so the findings from these studies should be taken as tentative and replicated to ensure validation. Also, since the normative data provided for these studies appear to lack African American participation, we must be careful to generalize these findings to this cultural group. Ideally, these studies would be redone with a heavier emphasis on African American participation.

CES-D

The Lewinsohn, Rohde, and Seeley (1996) CES-D screening items were discovered to have some predictive validity for later suicidal behavior. Roughly, 16.7 % of adolescents who were assessed as having high ideation at an initial screening made a suicide attempt within the subsequent year. Also, about 6.7 % of adolescent reporting moderate ideation at an initial screening made suicide attempt within the next year. Of those with mild ideation at an initial screening, 2.8 % made suicide attempts within the following year, and of those who reported no suicidal ideation, only 0.3 % made suicide attempts in the next year.

Some of the studies reviewed in this section had samples that were overrepresented by males, calling into question the utility of the CES-D across genders. Additionally, although both supplementary suicide screening sets (Garrison et al. 1991; Lewinsohn 1996) have been used extensively, there is a legitimate dearth of research on their psychometric properties. This lack of knowledge on their reliability and validity is a cause for concern, given the extent to which both are actively used, so research on the CES-D should shift its focus to this area.

CDI On the CDI, suicidal ideation is measured with one item and has the following response choices: 0 = "I do not think about killing myself;" 1 = "I think about killing myself but would not do it;" or 2 = "I want to kill myself." A score of 1 or 2 on this item indicates the presence of suicidal thoughts. This item refers to thoughts about suicidal actions with "nonzero intent to die," which is the widely accepted definition for suicidal behavior (O'Carroll et al. 1996). Since the CDI does not include any items that directly assess suicide attempts, it is not deemed useful for evaluating suicide attempters (Goldston 2000).

CDRS Although, the CDRS is primarily meant for the assessment of depression, it also has two items revolving around suicide, suicide ideation, and morbid ideation. The Suicide and Suicide Ideation item is on a 6-point scale, where 0 = no information, 1 = none, 2 = has thoughts about suicide—usually when angry, 3 = recurrent thoughts of suicide, 4 = thinks about suicide and names methods or if depressed, strongly denies thinking about suicide, and 5 = suicide attempt within the last month or actively suicidal. The Morbid Ideation item is on a 5-point scale, where 0 = no information, 1 = none expressed, 2 = some morbid thoughts—all related to a recent reality event, 3 = admits to morbid thoughts on questioning but does not dwell on them, or parents report morbid thoughts of child, and 4 = death themes spontaneously discussed or elaborate and extensive morbid ideation (Poznanski et al. 1979; Poznanski and Mokros 1996).

DISC

The subsection of the NIMH DISC 2.3 relevant for assessing suicide and depression focuses on respondent's thoughts of death, suicidal ideation, a specified suicide plan, and the association of suicidal thoughts with dysphoric symptoms. These inquiries reference the 2 weeks and 6 months preceding the interview (Goldston 2000; Shaffer et al. 2000). The DISC 2.3 also requests information from respondents regarding methods of suicide attempts, lifetime suicide attempts, number of suicide attempts, age at first suicide attempt, suicide attempts within the last 6 months, and suicide attempts within the context of dysphoria (Goldston 2000; Shaffer et al. 2000).

Conclusions

Not surprisingly, there is a paucity of research on the assessment of suicidality and depression in African American adults and youth. While we have elucidated the measures of which we are aware with any substantial psychometric evaluation with African Americans, we are limited in our discussions on this topic given that we could identify no measures with a specific focus on African Americans. New measures are currently being examined to assess suicide in multicultural populations (e.g., the Cultural Assessment of Risk for Suicide measure—CARS (Chu et al. 2013) but more research is needed to generate a significant group of measures readily identifiable as highly relevant to the needs of African Americans.

Review of Depression and Suicide Assessments for African American Adults and Youth

	Adults	
Assessment name	Disorder assessed	Recommendation(s) and/or relevant research findings
Beck Depression Inventory II	Major depressive disorder	Used with adults and youth; tested with African Americans; possibly efficacious
The Center for Epidemiology Studies Depression (CES-D) Scale	Major depressive disorder	Used with adults and youth; tested with African Americans; possibly efficacious
Patient Health Questionnaire—9 (PHQ-9)	Major depressive disorder	Primarily for adults; tested with African Americans; possibly efficacious
The Scale of Suicide Ideation (SSI)	Suicidality	Primarily for adults; tested with African Americans; possibly efficacious
Suicide Probability Scale	Suicidality	Primarily for adults; tested with African Americans; possibly efficacious
Subscales of measures		
Beck Depression Inventory Suicide Subscale	Suicidality	Primarily for adults; tested with African Americans; possibly efficacious
	Youth	
Children's Depression Rating Scale Revised	Depression	For youth; tested with African Americans Possibly efficacious, use with caution for cultural relevance
Reynold's Adolescent Depression Scale	Depression	
Children's Depression Inventory	Depression	For youth; very limited use with African Americans Use for screening, not diagnosis and follow up with culturally relevant depression diagnostic measure
Suicide Ideation Questionnaire-Jr	Suicidality	Probably efficacious with African American youth
DISC	Suicidality	Due to expense and training needed, this is not a practical measure for most clinicians to consider
Subscales of measures		
CES-D	Suicidality	Possibly efficacious
DISC	Suicidality	Possibly efficacious
CDRS	Suicidality	Possibly efficacious
Kiddie-Schedule for affective disorders and Schizophrenia	Suicidality	Cumbersome and possibly less useful for clinicians

References

Alegria, M., Chatterji, P., Wells, K., Cao, Z., Chen, C.-N., Takeuchi, D., Jackson, J., Meng, X.-L. (2008). Disparity in depression treatment among racial and ethnic minority populations in the United States. *Psychiatric Services,59*(11), 1264–1272. doi:10.1176/appi.ps.59.11.1264.

Allgaier, A.-K., Frühe, B., Pietsch, K., Saravo, B., Baethmann, M., & Schulte-Körne, G. (2012). Is the Children's Depression Inventory short version a valid screening tool in pediatric care? A comparison to its full-length version. *Journal of Psychosomatic Research,73*(5), 369–374. doi:http://dx.doi.org/10.1016/j.jpsychores.2012.08.016.

Arnau, R. C., Meagher, M. W., Norris, M. P., & Bramson, R. (2001). Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients (Article; proceedings paper). *Health Psychology,20*(2), 112–119. doi:10.1037//0278-6133.20.2.112.

Beck, A. T., & Steer, R. A. (1991). Manual for the Beck scale for suicide ideation. San Antonio: Psychological Corporation.

Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of suicidal intention—scale for suicide ideation. *Journal of Consulting and Clinical Psychology,47*(2), 343–352. doi:10.1037//0022-006x.47.2.343.

Beck, A. T., Steer, R. A., Kovacs, M., & Garrison, B. (1985). Hopelessness and eventual suicide—a 10-year prospective-study of patients hospitalized with suicidal ideation. *American Journal of Psychiatry,142*(5), 559–563.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review,8*(1), 77–100.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for the Beck Depression Inventory-II. San Antonio: Psychological Corporation.

Beck, A. T., Brown, G. K., & Steer, R. A. (1997). Psychometric characteristics of the scale for suicide ideation with psychiatric outpatients. *Behaviour Research and Therapy,35*(11), 1039–1046. doi:10.1016/S0005-7967(97)00073-9.

Breland-Noble, A. M., & AAKOMA Project Advisory Boards. (2012). Community and treatment engagement for depressed African American youth: The AAKOMA FLOA pilot. *Journal of Clinical Psychology in Medical Settings,19*(1), 41–8.

Breland-Noble, A. M., & Weller, B. (2012). Examining African American adolescent depression in a community sample: The impact of parent/child agreement. *Journal of Child and Family Studies,21*(5), 869–876. doi:10.1007/s10826-011-9547-z.

Breland-Noble, A. M., Bell, C., & Nicolas, G. (2006). Family first: The development of an evidence-based family intervention for increasing participation in psychiatric clinical care and research in depressed African American adolescents. *Family Process,45*(2), 153–169.

Breland-Noble, A. M., Burriss, A., & Poole, H. K. (2010). Engaging depressed African American adolescents in treatment: Lessons from the AAKOMA PROJECT. *Journal of Clinical Psychology,66*(8), 868–879.

Breland-Noble, A. M., Bell, C. C., Burriss, A., & AAKOMA Project Adult Advisory Board. (2011). Mama just won't accept this: Adult perspectives on engaging depressed African American teens in clinical research and treatment. *Journal of Clinical Psychology in Medical Settings,18*(3), 225–234. doi:10.1007/s10880-011-9235-6.

Brook, J. S., Stimmel, M. A., Zhang, C., & Brook, D. W. (2010). The association between earlier marijuana use and subsequent academic achievement and health problems: A longitudinal study. *The American Journal on Addictions,17*(2), 155–160.

Brown, G. K. (2002) *A review of suicide assessment measures for intervention research with adults and older adults*. Retrieved from http://www.suicidology.org/c/document_library/get_file?folderId=235&name=DLFE-113.pdf.

Brown, G. K., Beck, A. T., Steer, R. A., & Grisham, J. R. (2000). Risk factors for suicide in psychiatric outpatients: A 20-year prospective study. *Journal of Consulting and Clinical Psychology,68*(3), 371.

Bush, B. A., Novack, T. A., Schneider, J. J., & Madan, A. (2004). Depression following traumatic brain injury: The validity of the CES-D as a brief screening device. *Journal of Clinical Psychology in Medical Settings,11*(3), 195–201.

Campbell, M. H., Maynard, D., Roberti, J. W., & Emmanuel, M. K. (2012). A comparison of the psychometric strengths of the public-domain Zung self-rating depression scale with the proprietary beck Depression Inventory-II in Barbados. *West Indian Medical Journal,61*(5), 483–488.

Carey, M. P., Faulstich, M. E., Gresham, F. M., Ruggiero, L., & Enyart, P. (1987). Children's Depression Inventory: Construct and discriminant validity across clinical and nonreferred (control) populations. *Journal of Consulting and ClinicalPsychology,55*(5), 755–761. doi:10.1037/0022–006x.55.5.755.

Centers for Disease Control and Prevention (CDC), & National Center for Injury Prevention and Control. (2005). Web-Based Injury Statistics Query and Reporting System (WISQARS). http://www.cdc.gov/injury/wisqars/index.html. Accessed 20 Mar 2013.

Chu, J., Floyd, R., Diep, H., Pardo, S., Goldblum, P., & Bongar, B. (2013). A tool for the culturally competent assessment of suicide: The cultural assessment of risk for suicide (CARS) measure. *PsychologicalAssessment, 25*(2), 424–434.

Comstock, G. W., & Helsing, K. J. (1977). Symptoms of depression in two communities. *Psychological Medicine,6*(4), 551–563. doi:10.1017/S0033291700018171.

Crosby, A. E., & Molock, S. D. (2006). Introduction: Suicidal behaviors in the African American community. *Journal of Black Psychology,32*(3), 253–261. doi:10.1177/0095798406290552.

Cull, J. G., & Gill, W. S. (1988). Suicide Assessment Scale Manual.

Finch, A. J., Jr., Saylor, C. F., Edwards, G. L., & McIntosh, J. A. (1987). Children's Depression Inventory: Reliability over repeated administrations. *Journal of Clinical Child Psychology,16*(4), 339–341.

Frazier, T. W., Demeter, C. A., Youngstrom, E. A., Calabrese, J. R., Stansbrey, R. J., McNamara, N. K., & Findling, R. L. (2007). Evaluation and comparison of psychometric instruments for pediatric bipolar spectrum disorders in four age groups. *Journal of Child and Adolescent Psychopharmacology,17*(6), 853–866.

Garrison, C. Z., Addy, C. L., Jackson, K. L., & McKeown, R. E. (1991). A longitudinal study of suicidal ideation in young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry,30*(4), 597–603. Accessed 20 Mar 2013.

Goldston, D. B. (2000). *Assessment of suicidal behaviors and risk among children and adolescents*. Technical report submitted to the National Institute of Mental Health under Contract No. 263-MD-909995.

Goldston, D. B., Molock, S. D., Whitbeck, L. B., Murakami, J. L., Zayas, L. H., & Hall, G. C. N. (2008). Cultural considerations in adolescent suicide prevention and psychosocial treatment. *American Psychologist,63*(1), 14–31. doi:10.1037/0003-066x.63.1.14.

Gould, M. S., King, R., Greenwald, S., Fisher, P., Schwab-Stone, M., Kramer, R., Goodman, S., Canino, G., Shaffer, D. (1998). Psychopathology associated with suicidal ideation and attempts among children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry,37*(9), 915–923. doi:10.1097/00004583-199809000-00011.

Grothe, K. B., Dutton, G. R., Jones, G. N., Bodenlos, J., Ancona, M., & Brantley, P. J. (2005). Validation of the Beck Depression Inventory-II in a low-income African American sample of medical outpatients. *Psychological Assessment,17*(1), 110–114. doi:10.1037/1040–3590.17.1.110.

Harris, T. L., & Molock, S. D. (2000). Cultural orientation, family cohesion, and family support in suicide ideation and depression among African American college students. *Suicide Life Threat Behavior,30*(4), 341–353.

Hodges, K. (1990). Depression and anxiety in children: A comparison of self-report questionnaires to clinical interview. *Psychological Assessment: AJournal of Consulting and Clinical Psychology,2*(4), 376–381. doi:10.1037/1040-3590.2.4.376.

Horwood, L. J., Fergusson, D. M., Hayatbakhsh, M. R., Najman, J. M., Coffey, C., Patton, G. C., Silins, E., Hutchinson, D. M. (2010). Cannabis use and educational achievement: Findings from three Australasian cohort studies. *Drug and Alcohol Dependence,110*(3), 247–253. doi:10.1016/j.drugalcdep.2010.03.008.

Joe, S. (2006). Implications of national suicide trends for social work practice with Black youth. *Child and Adolescent Social Work Journal,23*(4), 458–471. doi:10.1007/s10560-006-0064-7.

Joe, S., Baser, R. E., Breeden, G., Neighbors, H. W., & Jackson, J. S. (2006). Prevalence of and risk factors for lifetime suicide attempts among blacks in the United States. *Journal of the American Medical Association,296*(17), 2112–2123.

Joe, S., Woolley, M. E., Brown, G. K., Ghahramanlou-Holloway, M., & Beck, A. T. (2008). Psychometric properties of the Beck Depression Inventory-II in low-income, African American suicide attempters. *Journal of Personality Assessment,90*(5), 521–523.

King, C. A., Segal, H., Kaminski, K., Naylor, M. W., Ghaziuddin, N., & Radpour, L. (1995). A prospective study of adolescent suicidal behavior following hospitalization. *Suicide and Life-Threatening Behavior,25*(3), 327–338.

Kneipp, S. M., Kairalla, J. A., Stacciarini, J., & Pereira, D. (2009). The Beck Depression Inventory-II factor structure among low-income women. *Nursing Research,58*(6), 400–409.

Kovacs, M. (1985). The children's depression, inventory (CDI). *Psychopharmacology Bulletin,21*(4), 995–998.

Kovacs, M. (2010). *Children's Depression Inventory 2 Manual*: Multi-Health Systems, Inc.

Krefetz, D. G., Steer, R. A., Gulab, N. A., & Beck, A. T. (2002). Convergent validity of the Beck Depression Inventory-II with the Reynolds adolescent depression scale in psychiatric inpatients. *Journal of Personality Assessment,78*(3), 451–460.

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression and diagnostic severity measure. *Psychiatric Annals,32,* 509–521.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine,16*(9), 606–613. doi:10.1046/j.1525-1497.2001.016009606.x.

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1993). Psychosocial characteristics of adolescents with a history of suicide attempt. *Journal of the American Academy of Child and Adolescent Psychiatry,32*(1), 60–68.

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1994). Psychosocial risk factors for future adolescent suicide attempts. *Journal of Consulting and Clinical Psychology,62*(2), 297.

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1996). Adolescent suicidal ideation and attempts: Prevalence, risk factors, and clinical implications. *Clinical Psychology: Science and Practice,3*(1), 25–46.

Maier, K. J., Goble, L. A., Neumann, S. A., Giggey, P. P., Suarez, E. C., & Waldstein, S. R. (2009). Dimensions across measures of dispositional hostility, expressive style and depression show some variation by race/ethnicity and gender in young adults. *Journal of Social and Clinical Psychology,28*(10), 1199–1225.

Mayes, T. L., Bernstein, I. H., Haley, C. L., Kennard, B. D., & Emslie, G. J. (2010). Psychometric properties of the children's depression rating scale-revised in adolescents. *Journal of Child Adolescent Psychopharmacology,20*(6), 513–516. doi:10.1089/cap.2010.0063.

McLaughlin, K. A., Hilt, L. M., & Nolen-Hoeksema, S. (2007). Racial/ethnic differences in internalizing and externalizing symptoms in adolescents. *Journal of Abnormal Child Psychology,35*(5), 801–816. doi:10.1007/s10802-007-9128-1.

Miranda, J., Siddique, J., Belin, T. R., & Kohn-Wood, L. P. (2005). Depression prevalence in disadvantaged young black women. *Social Psychiatry and Psychiatric Epidemiology,40*(4), 253–258.

Molock, S. D., Puri, R., Matlin, S., & Barksdale, C. (2006). Relationship between religious coping and suicidal behaviors among African American adolescents. *Journal of Black Psychology,32*(3), 366–389.

Multi-Health Systems Inc. (2011). CDI-2 Info Sheet USA 2011. Accessed 12 June 2011.

Nelson, W. M., Politano, P. M., Finch, A., Wendel, N., & Mayhall, C. (1987). Children's Depression Inventory: Normative data and utility with emotionally disturbed children. *Journal of the American Academy of Child and Adolescent Psychiatry,26*(1), 43–48.

Nguyen, H. T., Kitner-Triolo, M., Evans, M. K., & Zonderman, A. B. (2004). Factorial invariance of the CES-D in low socioeconomic status African Americans compared with a nationally representative sample. *Psychiatry Research,126*(2), 177–187. doi:http://dx.doi.org/10.1016/j.psychres.2004.02.004.

O'Carroll, P. W., Berman, A. L., Maris, R. W., Moscicki, E. K., Tanney, B. L., & Silverman, M. M. (1996). Beyond the tower of Babel: A nomenclature for suicidology. *Suicide and Life-Threatening Behavior,26*(3), 237–252.

Okwumabua, J. O., Wong, S. P., & Duryea, E. J. (2003). Depressive symptoms and decision making among African American youth. *Journal of Adolescent Research,18*(5), 436–453. doi:10.1177/0743558403255062.

Orvaschel, H., & Puig-Antich, J. (1987). Schedule for affective disorder and schizophrenia for school-age children: Epidemiologic version: Kiddie-SADS-E (K-SADS-E).

Osman, A., Gutierrez, P., Bagge, C. L., Fang, Q., & Emmerich, A. (2010). Reynolds adolescent depression scale-second edition: A reliable and useful instrument. *Journal of Clinical Psychology,66*(12), 1324–1345.

Pinto, A., Whisman, M. A., & McCoy, K. J. M. (1997). Suicidal ideation in adolescents: Psychometric properties of the suicidal ideation questionnaire in a clinical sample. *Psychological Assessment,9*(1), 63.

Poznanski, E. O., & Mokros, H. B. (1996). Children's depression rating scale, revised (CDRS-R): Manual: Western Psychological Services.

Poznanski, E. O., Cook, S. C., & Carroll, B. J. (1979). A depression rating scale for children. *Pediatrics,64*(4), 442–450.

Poznanski, E. O., Grossman, J. A., Buchsbaum, Y., Banegas, M., Freeman, L., & Gibbons, R. (1984). Preliminary studies of the reliability and validity of the childrens depression rating-scale. *Journal of the American Academy of Child and Adolescent Psychiatry,23*(2), 191–197. doi:10.1097/00004583-198403000-00011.

Poznanski, E. O., Freeman, L. N., & Mokros, H. B. (1985). Children's depression rating-scale—revised (September 1984). *Psychopharmacology Bulletin,21*(4), 979–989.

Radloff, L. S. (1977). The CES-D scale a self-report depression scale for research in the general population. *Applied Psychological Measurement,1*(3), 385–401.

Ramos, B. M., Carlson, B. E., & McNutt, L. A. (2004). Lifetime abuse, mental health, and African American women. *Journal of Family Violence,19*(3), 153–164. doi:10.1023/B:JOFV.0000028075.94410.85.

Renaud, J., Brent, D. A., Birmaher, B., Chiappetta, L., & Bridge, J. (1999). Suicide in adolescents with disruptive disorders. *Journal of the American Academy of Child andAdolescent Psychiatry,38*(7), 846–851.

Reynolds, W. M. (1986). Assessment of depression in adolescents: Manual for the Reynolds adolescent depression scale. Odessa: Psychological assessment resources.

Reynolds, W. M. (1987). Professional manual for the suicidal ideation questionnaire. Lutz: Psychological Assessment Resources.

Reynolds, W. M. (2002). Reynolds adolescent depression scale: RADS-2: Professional manual: Psychological Assessment Resources.

Reynolds, W. M., & Mazza, J. J. (1999). Assessment of suicidal ideation in inner-city children and young adolescents: Reliability and validity of the suicide ideation questionnaire-JR. *School Psychology Review,28*(1), 17–30.

Rohde, P., Mace, D. E., & Seeley, J. R. (1997). The association of psychiatric disorders with suicide attempts in a juvenile delinquent sample. *Criminal Behaviour and Mental Health,7*(3), 187–200.

Rozario, P. A., & Menon, N. (2010). An examination of the measurement adequacy of the CES-D among African American women family caregivers. *Psychiatry Research,179*(1), 107–112. doi:10.1016/j.psychres.2010.06.022.

Sashidharan, T., Pawlow, L. A., & Pettibone, J. C. (2012). An examination of racial bias in the Beck Depression Inventory-II. *Cultural Diversity and Ethnic Minority Psychology,18*(2), 203–209. doi:10.1037/a0027689.

Saylor, C. F., Finch, A. J., Spirito, A., & Bennett, B. (1984). The Children's Depression Inventory: A systematic evaluation of psychometric properties. *Journal of Consulting and Clinical Psychology,52*(6), 955.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH diagnostic interview schedule for children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry,39*(1), 28–38. doi:10.1097/00004583-200001000-00014.

Snowden, L. R., & Pingitore, D. (2002). Frequency and scope of mental health service delivery to African Americans in primary care. *Mental Health Services Research,4*(3), 123–130.

Spirito, A., Stark, L., Fristad, M., Hart, K., & Owens-Stively, J. (1987). Adolescent suicide attempters hospitalized on a pediatric unit. *Journal of Pediatric Psychology,12*(2), 171–189.

Spitzer, R. L., Williams, J. W., Kroenke, K., Linzer, M., deGury, F. V., Hahn, S. R., Brody, D., Johnson, J. G. (1994). Utility of a new procedure for diagnosing mental disorders in primary care: The prime-md 1000 study. *Journal of the American Medical Association,272*(22), 1749–1756. doi:10.1001/jama.1994.03520220043029.

Spitzer, R. L., Kroenke, K., Williams, J. W., & The Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of prime-md: The phq primary care study. *Journal of the American Medical Association,282*(18), 1737–1744. doi:10.1001/jama.282.18.1737.

Spitzer, R. L., Williams, J. B. W., Kroenke, K., Hornyak, R., & McMurray, J. (2000). Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD patient health questionnaire obstetrics-gynecology study. *American Journal of Obstetrics and Gynecology,183*(3), 759–769.

Steer, R. A., Ball, R., Raneeri, W. F., & Beck, A. T. (1997). Further evidence for the construct validity of the Beck Depression Inventory-II with psychiatric outpatients. *Psychological Reports,80*(2), 443–446.

Tamburrino, M. B., Lynch, D. J., Nagel, R. W., & Smith, M. K. (2009). Primary care evaluation of mental disorders (PRIME-MD) screening for minor depressive disorder in primary care. *Primary Care Companion to the Journal of Clinical Psychiatry,11*(6), 339.

U.S. Census Bureau. (2010). State and county quick facts. http://quickfacts.census.gov/qfd/states/00000.html. Accessed 20 Mar 2013.

U.S. Department of Health and Human Services, P. H. S., Office of the Surgeon General. (2001). Mental health: Culture, race, and ethnicity—A supplement to mental health: A report of the surgeon general (pp. 203). Rockville: U.S. Department of Health and Human Services.

Williams, C. D., Taylor, T. R., Makambi, K., Harrell, J., Palmer, J. R., Rosenberg, L., & Adams-Campbell, L. L. (2007). CES-D four-factor structure is confirmed, but not invariant, in a large cohort of African American women. *Psychiatry Research,150*(2), 173–180. doi:http://dx.doi.org/10.1016/j.psychres.2006.02.007.

Assessing the African American Child and Adolescent: Special Considerations and Assessment of Behavioral Disorders

8

Shawn C. T. Jones, Enrique W. Neblett, Ashly L. Gaskin
and Daniel B. Lee

The historic and disproportionate burden of disadvantages associated with social stratification due to racism, poverty, and unequal access to high quality education places African American children and adolescents at increased risk for the development of behavioral disorders (Hill et al. 2011). African American children—who account for nearly 27 % of the US Black population—report frequent experiences of discrimination (Neblett et al. 2008; Seaton et al. 2008), have the greatest percentage living in poverty of any ethnic group (38.2 %), and are found to be at increased risk for lack of educational readiness, underperformance, and school failure. Moreover, some studies indicate that African American youth have higher rates of conduct and attention problems than European American youth (e.g., Bird et al. 2001; Last and Perrin 1993). In light of evidence highlighting racial discrimination, poverty and its correlates (e.g., disorganized neighborhoods, parental stress, and exposure to violence), and academic failure as precursors to mental health and conduct problems (Hill et al. 2011), African American children and adolescents represent an essential population in which to examine the assessment of behavioral disorders.

The nosology informing the assessment of mental health of children and adolescents is the *Diagnostic and Statistical Manual, 4th edition, text revision* (DSM-IV-TR; American Psychiatric Association 2000). The main behavioral disorders outlined among children are Attention Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD). Prevalence rates of ADHD among African American children (based on parent or child report) have ranged from 2.1 % (Angold et al. 2002) to 5.7 % (Roberts et al. 2006). For ODD and CD, prevalence rates have been reported as 1.1 % and 5.5 %, respectively (Angold et al. 2002). In these studies, African American youth have reported lower rates of these disruptive behavioral disorders (DBDs) than their White counterparts. Thus, it appears that despite the aforementioned sociocultural risk factors, African American children are exhibiting resilience as it pertains to psychological adjustment. Nevertheless, findings related to the prevalence of DBDs have varied greatly depending on the reporter, making the assessment of behavioral disorders in African American children a vital undertaking.

This chapter will examine some of the most common methods for assessing behavioral disorders in African American youth. While it is

---

S. C. T. Jones (✉) · E. W. Neblett · A. L. Gaskin ·
D. B. Lee
Department of Psychology, University of North Carolina
at Chapel Hill, 243E Davie Hall, Chapel Hill,
NC 27599, USA
e-mail: jonessc@email.unc.edu

E. W. Neblett
e-mail: eneblett@email.unc.edu

A. L. Gaskin
e-mail: algaskin@email.unc.edu

D. B. Lee
e-mail: leedanb@email.unc.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_8,

acknowledged that the assessment of these disorders can be accomplished using interview methods, behavioral observation, and rating scales, this chapter focuses on the latter. Behavior rating scales provide a cost-effective approach for assessing behavioral problems and have been developed to cover a broad range of problems, behaviors, as well as strengths. Before reviewing these assessment tools, we begin the chapter with a brief discussion of cultural factors that should be considered in the assessment of behavioral disorders.

Cultural Considerations

Researchers have lamented the overrepresentation of deficit-oriented perspectives that have characterized the study of African American and other racial and ethnic minority children and youth, and called for a greater focus on positive youth development and outcomes (Cabrera et al. 2012; García Coll et al. 1996; Lambert et al. 2001). Deficit-oriented approaches emphasize elevated rates of behavioral problems and fail to incorporate culturally-informed perspectives on African American children and adolescents. The accurate assessment of both the strengths of and psychopathology in African American youth, however, necessitates consideration of the historical context of African American youth, acknowledgement of racism, and other societal ills with which African American youth contend daily (e.g., poor living conditions, unsafe neighborhood, and high stress environments), and thoughtful reflection about how these challenges may inform the presentation of African American youths' problem behaviors. Two critical cultural considerations that must be taken into account when assessing behavioral disorders in African American children and adolescents are: (1) respondent (informant) bias and (2) the cultural appropriateness of assessment instruments. Addressing these areas will be critical in the accurate assessment of behavioral disorders in African American children, particularly in light of the cost of behavioral problems for academic failure and subsequent mental health and wellbeing (Hill et al. 2011).

Respondent Bias in Assessment of Behavioral Disorders

One of the most troubling concerns in the assessment of behavioral disorders for African American youth is the persistent discrepancy between prevalence rates of behavioral disorders reported by parents and children on the one hand, and teachers on the other (Worrell 2009). There are few large-scale epidemiological studies of youth mental disorders, but several studies find lower or comparable rates of attention-deficit and disruptive behavioral disorders for African American youth as compared to White youth on the basis of parent and child reports (Angold et al. 2002; Pastor and Reuben 2005; Roberts et al. 2006; see Cuffe et al. 2005 for a notable exception). Measures based on teacher ratings, however, overwhelmingly find higher rates of disruptive behavioral disorder for African American youth (Hosterman et al. 2008; McDermott et al. 1993; Nolan et al. 2001; Reid et al. 1998; Worrell 2009). For instance, McDermott and Spencer (1997) reported higher than expected "numbers of African American youth exhibiting impulsive aggression and opposition defiance (about 1.5 times expectancy)" (p. 396) on the basis of standardized teacher observational scales. One leading hypothesis is that teacher bias may account for discrepancies in parent and child reports versus teacher ratings. According to this view, some teachers may not be fully aware of African American children's culture and environment and may report the behavior of African American youth as problematic when the child's culture defines the very same behavior as normative or even positive. Scholars have emphasized the normative role of verve and rhythm (i.e., rhythmic and creative movement, posture, speech patterns, and behavior as well as a preference for high stimulation environments) in African American culture (Cunningham and Boykin 2004), pointing out that the traditional didactic "teacher talks and students listen" mode of learning may be inconsistent with African American patterns of socialization. Thus, a child who exhibits high but culturally appropriate movement expressiveness or who exhibits a preference for high classroom stimulation consistent with how he or she has been socialized prior to school entry may be assessed as hyperactive, disruptive, or off task. Furthermore, aggressive or hostile behavior may actually be a normative or understandable response due to the challenges of coping with racial discrimination but may be misinterpreted by teachers as behavioral problems in a system where compliance and obedience are valued and emphasized (Lambert et al. 2001).

There is some empirical evidence of teacher biases and misinterpretation of behavior as a result of different cultural norms (Lambert et al. 2001; Worrell 2009); however, recent work in this area appears reluctant to embrace the possibility that biases may account for rater discrepancies (e.g., Epstein et al. 2005; Hosterman et al. 2008; Miller et al. 2009). Consistent with the view that teachers (and psychologists, see Mann et al. 1992) may report behavior of a child from another culture as problematic (see also Puig et al. 1999), Zimmerman et al. (1995) found that Caucasian and Hispanic teachers were more likely than African American teachers to endorse higher problem scores for African American youth. Moreover, the greatest discrepancy in scores was between African American parents' ratings and non-African American teachers. In terms of less clear findings, Neal et al. (2003) found that teachers rated both Black and White students exhibiting behavior consistent with African American verve and rhythm as lower in achievement and higher in aggression and that the teachers would be more likely to refer such students to be assessed for special education. African American students who did not exhibit this pattern of movement and rhythm, on the other hand, were rated as higher in achievement. In a recent review of ADHD in African Americans, Miller et al. (2009) concluded that parents' beliefs about ADHD, higher rates of risk and lack of treatment access and utilization may be more explanatory than teacher rating bias in understanding ratings differences, but acknowledged that the cause of such differences is far from settled. Indeed, few studies have rigorously examined the issue of informant bias and more studies will be needed to assess its role in the disparities between teachers' and others' ratings of behavioral problems and externalizing symptoms.

Appropriateness of Assessment Instruments

A second critical cultural consideration in the assessment of behavioral disorders is the cultural appropriateness of assessment instruments. On this score, scholars have raised concerns with respect to the appropriate norming of measures, as well as the content and cultural validity of rating scales, and the related domains of measurement equivalence of existing measures for African American and non-African American youth. With respect to appropriate norms and the standardization of available assessment instruments, Lambert et al. (2001) and others (e.g., Miller et al. 2009) have suggested that while test developers have tried to include representative samples of Blacks and other minorities in their studies, the children in these studies may not adequately represent the diversity of Black children and adolescents. Youth in the standardization sample do not necessarily represent the socioeconomic diversity of Black youth and also are typically drawn from traditional treatment facilities (e.g., community mental health clinics, hospitals, and special education programs) that do not include youth "triaged" to alternative facilities such as the child welfare and juvenile justice systems (Lambert et al. 2001).

In addition to measures that may not be appropriately normed, scholars have raised concerns about the content and cultural validity of rating instruments for behavioral disorders. It may be that particular items on assessment scales are culturally bound and do not reflect the cultural context or reality of African Americans and other minorities. Lambert et al. (2002) found that several problem behaviors African American children's parents reported to clinicians during intake interviews were not represented by items on widely used assessment instruments such as the Child Behavior Checklist (CBCL; reviewed later in this chapter), but instead, were similar to problems children of African descent in Jamaica reported to clinicians during intake interviews (see also Lambert et al. 1998). Interestingly, a small, but important group of studies finds evidence for and against construct equivalence for ADHD and externalizing rating scales for African American and Caucasian youth with the particular measure in question and the gender of the child seeming to account for differences in findings (see Epstein et al. 1998; Reid et al. 1998, 2001). Scholars have contended that the lack of metric equivalence in some cases may be explained by the meaning of words on some measures varying across groups or the measures themselves failing to adequately reflect common idiomatic expressions used by minority respondents (Lambert et al. 2001). Seminal work by Robert Williams (1972) has highlighted how the failure of items on a test to reflect the cultural context of the respondents can lead to inappropriate conclusions. Moreover, features of Black English such as semantic inversion—when the meaning of a word in Black English is the opposite of what it is in standard English (e.g., "bad" refers to "very good"; Smitherman 2004)—may lead to problems with the validity of available behavioral disorders assessments for African American children and youth.

While our discussion of cultural considerations is far from exhaustive, the evidence suggests that informant bias may play a role in our understanding of discrepancies among informants in the assessment of behavioral disorders. Rigorous and systematic investigations of this issue are needed to advance our understanding of the role bias may play. Our brief review also highlights the cultural appropriateness of assessment instruments in evaluating the presence of behavioral disorders. The content and cultural validity of rating scales may be influenced by how well existing measures capture the behaviors raters are likely to report as well as by factors such as the language used in the measure or assessment. Scholarly attention to and systematic investigations of both these issues will be critical in moving the field toward an accurate, culturally-informed, evidence-based approach to the assessment of behavioral disorders in African American children and adolescents.

Common Tests in the Assessment of Behavioral Disorders

The remainder of this chapter will focus on common instruments, tests, and procedures associated with the assessment of behavioral disorders among African American youth. We will review mainstream measures, as well as a tool developed specifically for use with Black children. The development of prominent assessment tools will be described, the conventional psychometric properties outlined, the utility of each measure for African American youth assessed, and where appropriate, recommendations offered. A summary of these evaluation criteria can be found in Table 8.1.

Structured and Semi-Structured Interviews

Although structured and semi-structured interviews are covered in another chapter in this volume (see Chap. 3), it is important to note that both types of interviews are integral to the assessment of behavioral disorders. These assessments tools have been developed in line with the main nosologies for cognitive, behavioral, and emotional disorders: the American Psychiatric Association's *Diagnostic and Statistical Manual* (DSM) and the World Health Organization's International Classification of Diseases (ICD). As such, these tools can easily assess common disruptive behavioral disorders (i.e., ADHD, ODD, and CD). Briefly, the most common standardized diagnostic interviews relevant to the assessment of behavioral disorders in African American children are the Diagnostic Interview Schedule for Children (DISC; Shaffer et al. 2000), a structured diagnostic interview, and the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS; Ambrosini 2000), a semi-structured interview instrument.

Table 8.1 Behavior Rating Scales used with African American youth

Assessment name	Disorders assessed	Demonstrated psychometric properties (validity, reliability)	Findings pertaining to utility	Recommendations
Achenbach System of Empirically Based Assessment	CD, ODD, ADHD, other disorders (such as anxiety and depressive symptoms)	Good to high validity and reliability for all scales	Gross et al. 2006 Jastrowski Mano et al. 2009 Lambert et al. 2002, 2002	Most popular measure for assessment of problems with children with substantial research. However, mixed findings related to utility necessitate further research with African American children
Behavioral Assessment System for Children—Second Edition	CD, ODD, ADHD, other disorders (such as anxiety and depressive symptoms)	Good to high validity and reliability for main scales. No technical information available for auxiliary scales	None	Lack of findings indicates that research is needed to speak to equivalence of scales and utility for African American youth
Conners 3rd edition	ADHD primarily, but also CD, ODD, and other symptomatology	Adequate to high validity and reliability for all scales	Epstein et al. 1998 Hosterman et al. 2008	Utility seems fairly established, particularly for assessment of ADHD
Eyberg Child Behavior Inventory	Conduct Problems	Generally adequate to high reliability and validity	Gross et al. 2007	Care should be taken when considering cutoff scores given findings related to mean differences
Child Symptom Inventory	ADHD, CD, ODD, plus an additional ten DSM-IV childhood disorders	Good to high reliability and validity	None	Research with larger samples of African American youth is needed
Strengths and Difficulties Questionnaire	CD, ODD, ADHD, other disorders, prosocial behavior	Adequate to high validity and reliability for all scales	Ruchkin et al. 2008	Despite multiethnic studies, limited data available for US born African American youth. Additional research needed

Questionnaires/Rating Scales

Achenbach System of Empirically Based Assessment

The Achenbach System of Empirically Based Assessment (ASEBA) is a multiple informant behavioral rating system that can be used to assess emotional and behavioral problems in youth. The ASEBA is the most widely used rating system for youth problems (Zamarripa and Lerma 2013). The system consists of three forms for two age groups: Parents complete the Child Behavior Checklist (CBCL/1.5–5, CBCL/6–18), teachers or caregivers complete the teacher report form (TRF/6–18) or the caregiver–teacher report form (C-TRF), and children 11–18 complete the youth self-report (YSR). Subscales for the CBCL and YSR are grouped into competence, syndrome, and DSM-oriented scales, the latter of which is based on criteria from the DSM-IV (DSM-IV; American Psychiatric Association 1994; Flanagan 2004).

Development and Norms

The ASEBA measurement tools have seen numerous iterations, with current scales having been more recently refined and norms having been updated from previous versions. The developers of the ASEBA boast that extensive procedures were utilized to obtain a nationally representative normative sample. These samples contained 1753, 1057, and 2319 individuals for the CBCL/6–18, YSR, and TRF, respectively, of which 20%, 20%, and 14%, were African American. Of the samples factor analyzed to determine the problem scales, roughly 20% were African American (Achenbach and Rescorla 2000). However, the percentage of clinically referred African Americans is not discussed in the manual, a limitation that has been addressed by other researchers with regards to utility (Zamarripa and Lerma 2013).

Scales and cutoff scores

The relevant syndrome scales for assessment of behavioral disorders are: *Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior,* and *Aggressive Behavior*. These syndromes are grouped as Total Problems, Internalizing Problems, and Externalizing Problems. Scores on these rating scales are given as T-scores, ranging from 50–100. In terms of cut points, T-scores above 65 indicate the borderline clinical range, raising a level of suspicion about problem areas, while T-scores of 70 and above are in the clinical range. The DSM-oriented scales, which include *Affective Problems, Anxiety Problems, Somatic Problems, Attention Deficit/Hyperactivity Problems, Oppositional Defiant Problems*, and *Conduct Problems*, have the same cutoff scores. There are no differential cutoff scores by race/ethnicity (Achenbach and Rescorla 2000).

Validity

Content validity for the current set of checklists is based on prior research with the scales (Zamarripa and Lerma 2013). Specifically, the problem items were developed and defined based on feedback from teachers, parents, students, and mental health professionals and DSM-related items were supported by experts' judgment (Achenbach et al. 2008). Criterion-related validity for ASEBA comes from various types of analyses. For instance, multiple regression analyses indicate that 2–33% of the variance on individual scales (CBCL/6–18, YSR, and TRF) is accounted for by referral status (i.e., those who were referred had significantly higher scores than nonreferred children; Flanagan 2004). Categorical analyses (e.g., chi square and odds ratios) show similar patterns. Construct validity has been established through correlational analyses with similar instruments, most commonly the BASC and Conners' (discussed below) with correlations averaging 0.69 (BASC) and 0.81 (Conners; Achenbach et al. 2008).

Reliability

Internal consistency reliabilities (coefficient alphas) for syndromes are strong across the checklists, with the strongest reliability for TRF (0.85) then CBCL (0.83) and YSR (0.79; Achenbach et al. 2008). Test–retest reliability and 12-month stability are also strong. Given that the CBCL and TRF can be administered by multiple informants (e.g., caregivers, teachers), cross-informant reliability is also important to consider, and mean agreement is substantial (0.76 and 0.60, respectively; Flanagan 2004).

Utility

Although ASEBA incorporated demographically representative proportions of African Americans in their normative samples, the issue of utility is still important to consider. In Achenbach et al.'s (2008) most recent discussion of multicultural assessment using these measures, they note that the comparisons of scores, psychometrics, and correlates are available for thousands of respondents from diverse populations, including at least 20 different societies (including countries in Africa; Achenbach et al. 2008; Zamarripa and Lerma 2013). Moreover, they report that the factor structures and results are broadly comparable. At the same time, there are some aspects of the ASEBA scales that have noted potential differences for African American populations. For instance, the manual notes that African American children obtained significantly higher scores on several of the TRF (teacher reported) items, though only a small portion of these items showed significant effect sizes (Achenbach and Rescorla 2000). An assessment of differential item functioning (DIF) noted that there were some items on the CBCL 1.5–5 that showed racial/ethnic and income level differences (Gross et al. 2006).

Potentially, more problematic than differences for African Americans at the item-level is the issue of whether the problem scales outlined by Achenbach (and based largely on non-Latino White samples) are equivalent for African Americans. While the findings have been mixed, Michael Lambert et al. (2002, 2003), and Jastrowski Mano et al. (2009) raise questions about the measure equivalence of the ASEBA rating scales. Lambert and colleagues have conducted extensive research and record reviews of African American and Jamaican children and have noted that in many cases the proposed factor structure did not hold for children of African descent (Lambert et al. 2003). In addition, content validity seems to be less firmly established when examining purely African American samples (Lambert et al. 2002). Jastrowski Mano et al. (2009) found similar issues related to model fit as well as reliability (internal consistency), but found that certain problem domains had adequate convergent and criterion validity (Jastrowski Mano et al. 2009).

Strengths, Limitations, Special Considerations

One of the biggest strengths of the ASEBA scales is its long history of research and implementation with diverse populations (Zamarripa and Lerma 2013). As stated earlier, these scales are the most widely used rating scale for the initial assessment of behavioral problems among children. Logistically, the scales can be easily scored using a computer-scoring program, and because of the multiple scales, cross-informant comparisons can be made. Given the scale development research, ASEBA seems to be a reasonable option for use with African American children. However, the recent research findings related to differential functioning of the items and lack of measurement equivalence warrant further research to determine how best to interpret these problem scales for African American children. Moreover, given the findings about higher endorsement of problem behaviors from teacher ratings, it is highly recommended that assessors collect information regarding the child's behavior from as many sources (parents, self-report, observation) as possible.

Behavior Assessment System for Children—Second Edition

The Behavior Assessment System for Children—Second Edition (BASC-2) is a multimethod rating system designed to assist in the educational classification of emotional and behavioral disorders in children as well as to facilitate the design of treatment plans (Stein 2007). Like ASEBA, the BASC primarily utilizes parent, teacher, and self-report: Parent Rating Scale (PRS), Teacher Rating Scale (TRS), and student Self-Report of Personality (SRP). However, unlike ASEBA, the BASC can also obtain background information (Structured Developmental History (SDH) and behavioral observation in the classroom (Student Observation System (SOS)). Moreover, because the BASC covers ages 2–25, the rating scales are available at the Preschool, Child, Adolescent, and College levels. According to the manual, the BASC can also be utilized to assist in diagnosing DSM-IV disorders, as well as AID in other educational planning and service determinations (Stein).

Development and Norms

The BASC-2 is currently in its second edition, with the developers noting a series of improvements and additions to each of the rating scales (Reynolds and Kamphaus 2004). The updated norms for the BASC were collected in the early 2000s, with over 13,000 children included in the normative sample. The developers note that these norms closely reflect US population estimates for race and ethnicity (as well as gender and maternal education). In fact, the manual presents an impressive breakdown of African American students for each age range, for each form, and compares the percentages to the relevant US percentages at that time. In total, the African American sample comprises about 15.7 % of the norming sample. The clinical norms (drawn from 1779 (TRS), 1975 (PRS), and 1527 (SRP) 4–18 year olds) were not matched on ethnicity in the same manner. African American percentages were 17, 14.4, and 13.1 respectively (Reynolds and Kamphaus).

Scales and cutoff scores

The items on the BASC-2 TRS, PRS, and part of the SRP are presented on a 4-point response scale. The BASC-2 is divided into Composite, Clinical, Adaptive, and Content Scales. The Composite Scales include the *Behavioral Symptoms Index*, *Internalizing*, and *Externalizing* scales. The Clinical Scales include domains similar to the ASEBA (e.g., Hyperactivity, Aggression, Conduct Problems, and Withdrawal) as well as unique domains such as Learning Problems and Atypicality (Stein 2007). The Content Scales also include relevant behavioral correlates such as anger control, bullying, and negative emotionality (Rescorla 2009). T-scores and percentile ranks are used for the three main measures, at both the composite and individual scale level. Scores at or above $T=70$ (97th percentile) define the clinically significant range, while scores $T=60–69$ (>84th percentile) represent the at-risk range (Rescorla 2009).

Validity

The majority of data on validity for the BASC-2 is based on intercorrelations (construct and concurrent validity; Stein 2007). The TRS showed high correlations with ASEBA's TRF, with internalizing problems correlating 0.64–0.80 and externalizing problems ranging 0.75–0.85. The Conners' Teacher Rating Scale was also correlated highly with the TRS for most scales, except those concerning anxiety. In addition, correlations were higher for children as opposed to adolescents. For the PRS and SRP, similar correlational patterns have been identified. In addition, the SRP scales have been correlated with other popular measures of internalizing problems (e.g., Children's Depression Inventory, Revised Children's Manifest Anxiety Scale (RCMAS), Beck Depression Inventory-II; Stein 2007). Construct validity is assessed by examining the Clinical scale and Composite score profiles by age group for each of the three main BASC measures. The manual shows that these profile patterns are generally consistent between age groups and reflect expected behavioral strengths and weaknesses. An exception in profile patterns (identified by the authors of the manual) exists for mood disorders (bipolar or depression; Stein 2007).

Reliability

Internal consistency reliabilities (alpha coefficients) are high for all three measures for the general norm group, with composite scores ranging from the low-to-mid 0.90s for the TRS and the mid-to-high 0.80s for the SRP (Stein 2007). The clinical norms group has similarly high coefficients (though at times lower). The reliabilities for individual scale scores are much lower than those for the composite scores, and as such, it is recommended that composite scores be used. Six-week (TRS and PRS) and 3-week (SRP) test–retest reliabilities ranged from 0.64 among adolescents' ratings on the TRS to 0.93 for teacher ratings of younger children. In general, the TRS (0.86) and the PRS (0.84) yielded the highest median reliability for children, whereas the SRP (0.84) yielded the highest median reliability for college-aged youth (Stein).

Strengths, Limitations, Special Considerations Surrounding Utility

As mentioned earlier, the BASC is often used in both school and clinical settings, and has been purported to answer several referral questions for youth (Reynolds and Kamphaus 2004). However, there is a dearth of evidence as to whether this system has been validated for use in all the relevant arenas (e.g., IEP planning, manifestation determination; Stein 2007). In addition, though the BASC-2 includes developmental history and observational rating measures, there is no technical information for these assessment instruments (Stein 2007). The scales are also unique in that they provide ways to check against biased responses, rater error, and other threats to validity. The BASC-2 can also be easily scored, provides rubrics for multiple levels of comparison, and report templates that can be used for parent feedback (Stein).

Perhaps the biggest limitation and consideration with regard to this measure is that there has been no published empirical work to confirm the utility of the BASC-2 for African American children. Although inferences could be drawn given the system's relationship with other behavioral assessment tools that have had more research undertakings (most notably ASEBA), research should be conducted to address issues such as measurement equivalence. This is particularly relevant given the nonrepresentative nature of the clinical developmental samples. A recent review of assessment measures for Latino children (Zamarripa and Lerma 2013) noted a similar dearth of evidence and suggested the BASC be used with proper inclusion of cultural data. While this is a reasonable recommendation, the importance of research to confirm psychometric and cultural validity and reliability cannot be understated.

Conners 3rd Edition

The Conners 3rd Edition (Conners 3) rating scales are another set of multiple informant assessment tools, used for children between 6 and 18 years of age (Arffa 2010). The primary purpose of the Conners 3 is to serve as a comprehensive assessment of Attention Deficit Hyperactivity Disorders (ADHD); however, the Conners scales are also commonly used to assess other conditions comorbid with ADHD. The Conners system can also be used to aid in the special education eligibility process, as a treatment evaluation tool, and for screening purposes (Arffa). Informants include parents (3-P), teachers (3-T), and children ages 8–18 (3-SR), and there are multiple versions (full, short, ADHD Index Global Index) available.

Development and Norms

The Conners 3 is a revision of the Conners Rating Scales-Revised (Conners 1997). This iteration of the forms was undertaken to keep in line with the most recent diagnostic information, update normative data, and link findings to interventions (Arffa 2010). While the focus of the forms was ADHD, the author also developed instruments to measure broader behavioral and emotional problems (Conners comprehensive Behavior Rating Scales) and to measure ADHD in younger children (Conners Early Childhood Rating Scales). In addition, language was changed to reflect DSM-IV symptoms, and impairment items were added. The standardization sample for the Conners 3 consisted of 50 males and 50 females in each age group (6–18 years for Teacher and Parent forms and 8–18 years for Self-Report forms) from the general population with an ethnic and racial distribution closely matching 2000 US Census data (Conners 2008). For instance, African American respondents for the Conners Parent report comprised 15.1 %, while comprising 15.7 % of the US population at that time. An additional clinical population of 718 youth for parent and self-report (15.32 % African American) and 694 youth for Teacher report (15.27 % African American was sampled). The Conners 3 was co-normed with the Conners Comprehensive Behavior Rating Scales (Conners).

Scales and cutoff scores

Informants for the Conners answer questions on a scale from 0 (*Never or Seldom*) to 3 (*Very true or Very frequently*). The number of questions ranges from 99 for the 3-SR to 115 for the 3-T (Full version), around 40 items for the short form, and 10 items for the ADHD and Global Indices (Arffa 2010). The relevant content area scales include *Inattention, Hyperactivity, Impulsivity, Learning Problems, Executive Functioning, Conduct Problems, Oppositional Behavior, and Social Problems*. Empirical scales include ADHD, conduct disorder (CD), and oppositional defiant disorder (ODD). In addition, there are scales that address the reporter biases and inconsistency in reporting. Linear T-scores and empirical percentiles are used to determine levels of concern. T-scores greater than 60 indicate elevated scores (more problems than are typically reported), while T-scores 70 and above indicate clinical levels of problems.

Validity

Criterion validity for the Conners scales is based on the form's ability to discriminate between clinical and nonclinical populations. According to the manual, the Conners 3 can accurately distinguish between clinical and nonclinical groups, as well as between clinical groups with ADHD and other behavioral or learning disorders. Construct validity has been established through factor analytic techniques, as well as cross-informant comparisons. The mean parent to teacher correlation was 0.60, the mean parent to youth correlation was 0.56, and the mean teacher to youth correlation was 0.48. Evidence of convergent validity was established via correlations with earlier versions of the Conners, as well as the ASEBA forms, BASC-2, and Behavior Rating Inventory of Executive Functioning (Gioia et al. 2000).

Reliability

Internal consistency coefficients are 0.90 or above for parent and teacher scales and 0.85 or above for Self-Report scales. Adjusted test–retest reliability ranges from 0.82 to 0.98 for Parent scales except for Executive Functioning (0.72) and Peer/Family Relations (0.78), from 0.83 to 0.90 for Teacher scales except for Executive Functioning (0.78), and from 0.71 to 0.83 for the Self-Report scale (Arffa 2010).

Utility

As previously stated, the Conners 3 and Comprehensive Behavior scales were normed with representative samples that included nonnegligible proportions of African Americans (though this was not the case for clinical samples). The Conners 3 manual also provides discussions of differences by race/ethnicity for each measure. According to the manual the 3-P, 3-T, and 3-SR showed significant effects by race/ethnicity at the multivariate level. For the parent report, analyses generally revealed that Caucasian parents reported more problem behaviors than parents of other racial groups. However, African American parents reported greater executive functioning than did Hispanic parents. For teacher reports, African American youth were rated as having higher levels of defiance/aggression and ODD than White or Latino youth. For self-report, Caucasian youth again scored higher than other racial/ethnic groups on relevant problem scales. However, the manual notes that all of these effects are small, accounting for only small amounts of variance explained (Arffa). In addition to these findings, research has generally supported that the Conners has factor equivalence for African Americans (e.g., Epstein et al. 1998; Hosterman et al. 2008). Of note, these studies were conducted with earlier versions of the Conners and were limited to the teacher rating forms. Moreover, the rationale for the mean differences and higher ratings by teachers has not been fully established.

Strengths, Limitations, and Special Considerations

Taken together, the Conners 3 and Comprehensive Behavior scales provide a broad means for assessing ADHD and comorbid disorders. The Conners is unique in its ability to differentiate subtypes of ADHD in a reliable manner. In terms of logistics, the various forms can be easily completed and scored online, and interpreted with relative ease. The near ubiquitous nature of the Conners rating scales in school and clinical settings makes them a reasonable option for the assessment of behavioral disorders (especially ADHD) in African American samples. Moreover, unlike the clearly stated issues with the ASEBA or the lack of findings for the BASC-2, initial research seems to indicate reasonable structural and measurement equivalence for the Conners. Nevertheless, like many of the other rating systems, mean ethnic differences, particularly among the Teacher scales, necessitate additional research and the use of informant scales that are more established in terms of their utility.

Other Mainstream Behavior Rating Scales: ECBI, CSI, SDQ

There are several additional rating scales that have been routinely used for the assessment of behavioral disorders. While these scales or rating systems have all shown acceptable validity and reliability in a general population, and many have norming samples that include African American reporters, their clinical utility for African American children has not been well established. As such, for each of these measures it is recommended that care be taken when using these scales, and where possible, that additional research regarding utility be undertaken. These scales are each described briefly below.

Eyberg Child Behavior Inventory

The Eyberg Child Behavior Inventory (ECBI) and Sutter-Eyberg Student Behavior Inventory Revised (SESBI-R) are self-report measures that assess conduct problems for children aged 2–16 years (Meikamp 2003). According to the manual, the ECBI assesses both the severity of behavior problems (intensity scale) and the extent to which parents and teachers find these behaviors problematic (problem scale), with the latter speaking to issues of impaired functioning. As such, these scales are often used to identify children in need of treatment for conduct problems. Each scale consists of approximately 40 items, with items either scored on a 7-point or Yes–No scale. T scores greater than 60 are noted as "Exceeds Cutoff," and warrant additional evaluation. The ECBI was normed using a reportedly representative sample in terms of ethnicity (including 19 % African Americans); however, the sample was based completely in the Southeastern US. Additionally, the lone standardization study lacked enough representativeness for meaningful comparisons by ethnicity. Despite these limitations, the ECBI and SESBI have been shown to exhibit adequate psychometric properties, with internal consistency estimates in the mid 0.90s and convergent validity with scales on other rating systems such as the CBCL ranging from 0.41–0.75 (Meikamp). In terms of the utility of the ECBI for African American youth, a recent study using a sample of Black (26.7 %) and Latino youth indicated that while reliabilities and construct validity (using CFA) were similar across ethnic groups, mean score differences were significant, suggesting lower ratings of intensity for African American children (Gross et al. 2007). The researchers stated that these lower ratings by African American parents could necessitate adjusting cutoff scores for identification of at-risk Black youth. Moreover, the ratings call into question the utility of the ECBI as a measure of treatment progress for African American youth. As this was the first study of its kind, more research is needed to verify these findings.

Child Symptom Inventory

The Child Symptom Inventory 4 (CSI-4) is a rating scale that screens for the behavioral and affective symptoms for thirteen common DSM-IV childhood disorders for children aged 5–12 years (Hoff 2005). The CSI was designed for use in clinical settings and provides a briefer alternative to structured interviews (see Chap. 3), and can denote whether additional assessment is needed (Hoff). Informants for the CSI can be either parents or teachers, and the CSI provides extensive overviews related to the diagnostic utility of items. Items for this measure are rated on a 4-point scale, and the CSI can be scored using both a categorical (symptom count) and dimensional (symptom severity) metric. Symptom count is scored using a 2-point scale, with cutoff scores related to receiving positive scores equal to or exceeding the minimum number of symptoms necessary for a DSM-IV diagnosis. Symptom severity compares gender-normed samples and calculates T-scores using a standard convention (i.e., $T > 70$ connotes severity). The psychometric properties for the CSI are well established. Internal consistency is adequate to good (range = 0.74–0.94 parent and 0.70–0.96 teacher) for all scales, and test–retest and interrater reliabilities are also adequate. Not surprisingly, criterion validity was established comparing the CSI to psychiatric diagnoses, and metrics (sensitivity and specificity) were moderate to high. Convergent and divergent validity have also been established using other common checklist and rating systems. The biggest threat to utility is the measure's norming samples: The most recent norms for parent ratings are based on 551 children, only 5 % of whom are African American. More concerning are the teacher norms, which are based on a 95 % Caucasian sample (Hoff). Given the potential utility of this measure for mapping onto DSM-IV diagnoses, interested practitioners should collect data specific to African Americans when possible.

Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire is a dimensional assessment instrument used to measure children's (3–16 years) problems and positive characteristics (Achenbach et al. 2008). Used primarily as a brief screening measure, the SDQ is based on the Rutter (1967) parent questionnaire, and consists of five items each for five difficulty scales: *Hyperactivity*, *Emotional Symptoms*, *Conduct Problems*, *Peer Problems*, and *Prosocial Behavior,* as well as a

Total Difficulties score. These items are based on nosological concepts from the DSM-IV and ICD-10 (Bourdon et al. 2005). Like the ASEBA, the SDQ is rated on a scale range from 0–2, with reverse scoring for the prosocial items. The SDQ is available in both parent and teacher forms, and a self-report version is available for adolescents. In addition to the 25-item scale, the SDQ can also include an impact supplement that can assess various indices related to functioning. The scale was originally tested on a sample of 403 children in London, and as such did not include data on race and ethnicity breakdown (Goodman 1997). Moreover, although the SDQ was included in a 2001 supplement to the National Health Interview Survey (NHIS) and administered to nearly 10,000 American parents, the accompanying report does not include a breakdown by race/ethnicity. Nevertheless, extensive psychometric properties have been established for the SDQ. Content validity is based on a comparison of the scales from the SDQ with the Rutter parent and teacher questionnaires. Correlations between the two range from 0.78 to 0.92 (Goodman 1997). Criterion validity for the SDQ has been established from a series of analyses discriminating between children with and without mental health problems. The primary analytic approach for the SDQ has been Receiver Operating Characteristics (ROC) analyses, which have indicated that the measure does an adequate job with regards to sensitivity and specificity, with an area under the curve (AUC) of 0.83 across scales (Achenbach et al. 2008). Construct validity has been established through a comparison of the SDQ with the ASEBA scales and other comparable measures, SDQ and ICD-10 diagnoses, genetic studies, and factor analytic methods. With regard to reliability, alpha coefficients averages range from 0.59 to 0.78 for the difficulty scales, and 0.83 for the Total Difficulties scale. Test–retest reliabilities have been reported in the low 0.70s for the difficulty scales, with reliabilities slightly better for the broad Total Difficulties scale (Achenbach et al. 2008).

The SDQ shares one of its biggest strengths as an assessment tool with the ASEBA scales, namely its use and application in a variety of cultures. In addition to the UK and USA, the SDQ has normative data available for other European nations and Australia, and findings are also published using the SDQ in several countries in Asia and Africa. Despite the robust multicultural findings internationally, findings supporting the utility of the SDQ forms with African American populations have been scant. In fact, one of the few published studies to use the SDQ with African American samples (57 % of 4671 children) suggested that the five factor structure proposed by Goodman may not provide the best fit in a multiethnic, urban US sample (Ruchkin et al. 2008). The authors suggest that further research be conducted to ensure utility, particularly for multiethnic and inner-city American youth. Thus, despite the SDQ's wide availability, multinational use, and usefulness as a quick screening measure, care should be taken when deciding to utilize this battery with African American youth.

Behavioral Assessment for Children of African Heritage (BACAH)

The Behavioral Assessment for Children of African Heritage (BACAH) is unique in that it is the only measure to date that has been designed specifically for Black children. The BACAH, developed by Michael Lambert et al. (2005) is also unique in that unlike the aforementioned rating scales, it was intentionally developed to assess behavioral and emotional *strengths*. In fact, the authors cite a point raised earlier in this chapter—namely that literature has generally adopted a deficit approach concerning African American children—as the impetus for the development of these scales. In addition, as has been shown with the other mainstream measures, there is a dearth of assessment research conducted with African American children, limiting the cultural validity of these measures and their utility for Black children. The BACAH is a set of forms "written in the voice of Black children" to be used for ages 4–16 years (Lambert et al. 2005). Similar to the mainstream rating scales, forms are available for parents, teachers, and youth self-report, and the measure also includes an interview schedule. The authors note that the forms can be used to scaffold clinical intervention, identify strengths that need reinforcing, aid in research with Black children, and evaluate treatment progress (Lambert et al. 2005).

Development and Norms

Given that the BACAH represented the first attempt to an emic approach to strengths based assessment for Black children, the authors note in the development article that great care was developed in identifying items and establishing norms. The development of the forms began with 20 focus groups that included African American parents, teachers, clinicians, and children (the clinician sample also included Latino and White professionals with expertise working with Black children). In addition, two distinct sets of 30 children, parents, and teachers were utilized to provide feedback on (a) the difficulty of the forms, (b) relevance of the items, and (c) clarity, as a means of ensuring content and cultural validity. The forms were then developed across informants in a manner that linked each of the rating scales to an identical metric. Methods employed to accomplish this included exploratory factor analysis (EFA) and item response theory (IRT).The development sample included 559 parents, 489 teachers, and 417 adolescents from 12 schools and 8 clinics in the Midwest and Northeast. Data were combined for referred and nonreferred children (Lambert et al. 2005).

Scales and Cutoffs

The BACAH forms each contain 64 items, with each item rated on a three point scale (*0 = not true, 1 = somewhat or sometimes true, 2 = very true or often true*). In addition, each strength item is rated as having a negative effect (1), no effect (0) or a positive effect (+1) on children. The BACAH consists of two scales: *Resilience* and *Self-Regulation and Prosocial Behavior*. Because the BACAH measures strengths, it does not have cutoff scores. Rather, the use of computerized adaptive testing (CAT) is used to guide areas of functioning (Lambert et al. 2005).

Psychometric Information: Validity and Reliability

As previously noted, great care was taken to ensure the content and cultural validity of the BACAH. Doing so was primarily accomplished through EFA and IRT. Specifically, the authors of the BACAH use IRT and Differential Item Functioning (DIF) analyses to test whether the items for the two factors loaded similarly across informant groups. Additionally, the authors tested item invariance across gender, SES, and referral status, as well as mean group differences by age group. Findings indicated that there was DIF only by referral status and across informants, and that DIF was only noted in a few items. In terms of reliability, coefficient alpha's indicated very good reliabilities for both resilience and self regulation and prosocial behavior for parents (0.96, 0.95) teachers (0.97, 0.96), and youth self-report (0.93, 0.96). Again, cross-informant development and analyses of items indicate that interrater reliability is also adequate (Lambert et al. 2005).

Strengths, Limitations, and Recommendations

The major strengths of the BACAH lie in it being one of the few published assessment tools for measuring behavioral and emotional strengths among Black children and the attention to detail related to the development of the various forms. This latter strength ensures that the measure is sensitive to many of the important factors for consideration outlined at the beginning of this chapter. Additionally, the use of computerized adaptive testing allows for a person-centered approach to assessing these strengths, while also minimizing biases such as practice effects. Despite these promising findings, there are also some limitations to the measure. First, because the measure was developed in such a unique manner, it was intentionally not compared to other mainstream forms (e.g., ASEBA). As such, there are unknowns about convergent and divergent validity. Additional research comparing the BACAH with other scales that measure strengths (such as the Behavioral and Emotional Rating Scale (BERS)) could do much to show the unique contribution of this measure. Second, a recent database search indicates that there is no additional published work utilizing the measure, a necessary, if not sufficient criteria for evidence-based assessment. The third limitation is only relevant given the focus of this chapter, namely that the BACAH by design does not measure behavioral problems, and as such cannot be used to aid in the diagnoses of clinical disorders. Thus, the measure's utility depends on the purposes of the clinician or researcher. However, even in cases where the primary referral question is diagnosis, the BACAH stands as a culturally relevant supplement, giving assessors invaluable information about the strengths and competencies of their African American clients.

Global Conclusions and Recommendations

As has been documented, there are several assessment tools—primarily rating scales—that can be used to assess behavioral disorders in African American children and adolescents. While it is clear from database searches that all of these measures are being used with African American youth and that these measures demonstrate traditional psychometric properties, it is notable that all of these measures are not created equal in terms of their demonstrated cultural validity and utility for this population. Recommendations for use have been provided for the three major rating systems discussed (i.e., ASEBA, BASC, Conners) and next steps for research have been outlined for the others. Clinicians are encouraged to keep in mind factors that may contribute to the presentation of symptoms that are being endorsed on these scales (e.g., discrimination, rater bias), and when possible to include an assessment of Black children's strengths and competences. Again, the BACAH provides an obvious choice, but many of the mainstream rating systems include supplements to measure hobbies, competencies, and prosocial behavior. In addition, given the issues surrounding informant biases (in particular higher mean endorsement of problems by teachers), direct observation of children and adolescent behavior on the part of the assessment team also may be prudent (e.g., Puig et al. 1999). As evidence based assessment approaches would dictate (e.g., Youngstrom 2008), choice of assessment tools should ultimately be driven by a given rating system's ability to predict, prescribe, or track process for African American youth, while fitting the unique needs and preferences of the population.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. Burlington: Research Center for Children, Youth, & Families, University of Vermont.

Achenbach, T. M., Becker, A., Dopfner, M., Heiervang, E., Roessner, V., Steinhausen, V. H., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry, 49*(3), 251–275.

Ambrosini, P. J. (2000). Historical development and present status of the schedule for affective disorders and schizophrenia for school-age children (K-SADS). *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(1), 49–58. doi:10.1097/00004583-200001000-00016.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev). Washington, DC: American Psychiatric Association.

Angold, A., Erkanli, A., Farmer, E. M. Z., Fairbank, J. A., Burns, B. J., Keeler, G., & Costello, E. J. (2002). Psychiatric disorder, impairment, and service use in rural African American and White youth. *Archives of General Psychiatry, 59*(10), 893–904. doi:10.1001/archpsyc.59.10.893.

Arffa, S. (2010). Review of the Conners 3rd edition. In *The eighteenth mental measurements yearbook*. http://www.unl.edu/buros/. Accessed 25 Mar 2013.

Bird, H. R., Canino, G. J., Davies, M., Zhang, H., Ramirez, R., & Lahey, B. B. (2001). Prevalence and correlates of antisocial behaviors among three ethnic groups. *Journal of Abnormal Child Psychology, 29*(6), 465–478. doi:10.1023/A:1012279707372.

Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The strengths and difficulties questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(6), 557–564. doi:10.1097/01.chi.0000159157.57075.c8.

Cabrera, N. J., Beeghly, M., & Eisenberg, N. (2012). Positive development of minority children: Introduction to the special issue. *Child Development Perspectives, 6*(3), 207–209. doi:10.1111/j.1750-8606.2012.00253.x.

Conners, C. K. (1997). *Conners' Rating Scales—revised technical manual*. Toronto: Multi-Health Systems.

Conners, C. K. (2008). *Conners' 3rd edition technical manual*. Toronto: Multi-Health Systems.

Cuffe, S. P., Moore, C. G., & McKeown, R. E. (2005). Prevalence and correlates of ADHD symptoms in the national health interview survey. *Journal of Attention Disorders, 9*(2), 392–401. doi:10.1177/1087054705280413.

Cunningham, R. T., & Boykin, A. W. (2004). Enhancing the cognitive (and school) performance of African American children: The infusion of Afro-cultural research-based practices. In R. L. Jones (Ed.), *Black psychology* (4th ed., pp. 487–508). Hampton: Cobb & Henry.

Epstein, J. N., March, J. S., Conners, C. K., & Jackson, D. L. (1998). Racial differences on the Conners Teacher Rating Scale. *Journal of Abnormal Child Psychology, 26*(2), 109–118. doi:10.1023/A:1022617821422.

Epstein, J. N., Willoughby, M., Valencia, E. Y., Tonev, S. T., Abikoff, H. B., Arnold, L. E., & Hinshaw, S. P. (2005). The role of children's ethnicity in the relationship between teacher ratings of attention-Deficit/Hyperactivity disorder and observed classroom behavior. *Journal of Consulting and Clinical Psychology, 73*(3), 424–434. doi:10.1037/0022-006X.73.3.424.

Flanagan, R. (2004). Review of the Achenbach system of empirically based assessment. In *The sixteenth mental measurements yearbook*. http://www.unl.edu/buros/. Accessed 25 Mar 2013.

García Coll, C., Crnic, K., Lamberty, G., & Wasik, B. H. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development, 67*(5), 1891–1914. doi:10.2307/1131600.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology, 6*(3), 235–238. doi:10.1076/chin.6.3.235.3152.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586. doi:10.1111/j.1469-7610.1997.tb01545.x.

Gross, D., Fogg, L., Young, M., Ridge, A., Cowell, J. M., Richardson, R., & Sivan, A. (2006). The equivalence of the child behavior Checklist/1 1/2-5 across parent race/ethnicity, income level, and language. *Psychological Assessment, 18*(3), 313–323. doi:10.1037/1040-3590.18.3.313.

Gross, D., Fogg, L., Young, M., Ridge, A., Cowell, J., Sivan, A., & Richardson, R. (2007). Reliability and validity of the Eyberg child behavior inventory with African-American and Latino parents of young children. *Research in Nursing & Health, 30*(2), 213–233. doi:10.1002/nur.20181.

Hill, N. E., Mann, T. L., & Fitzgerald, H. E. (2011). In Hill N. E., Mann T. L. & Fitzgerald H. E. (Eds.), *African American children and mental health, vols 1 and 2: Development and context, prevention and social policy*. Santa Barbara: Praeger/ABC–CLIO. Accessed 25 Mar 2013.

Hoff, K. E. (2005). Review of the Child Symptom Inventory-4. In *The sixteenth mental measurements yearbook*. http://www.unl.edu/buros/. Accessed 25 Mar 2013.

Hosterman, S. J., DuPaul, G. J., & Jitendra, A. K. (2008). Teacher ratings of ADHD symptoms in ethnic minority students: Bias or behavioral difference? *School Psychology Quarterly, 23*(3), 418–435. doi:10.1037/a0012668.

Jastrowski Mano, K. E., Davies, W. H., Klein-Tasman, B., & Adesso, V. J. (2009). Measurement equivalence of the child behavior checklist among parents of African American adolescents. *Journal of Child and Family Studies, 18*(5), 606–620. doi:10.1007/s10826-009-9263-0.

Lambert, M. C., Lyubansky, M., & Achenbach, T. M. (1998). Behavioral and emotional problems among adolescents of Jamaica and the United States: Parent, teacher, and self-reports for ages 12 to 18. *Journal of Emotional and Behavioral Disorders, 6*(3), 180–187. doi:10.1177/106342669800600306.

Lambert, M. C., Markle, F., & Bellas, V. F. (2001). Psychopathology and assessment in African American children and families: A historical, ecological, and strength-based perspective. In A. M. Neal-Barnett, J. M. Contreras, & K. A. Kerns (Eds.), *Forging links: African American children clinical developmental perspectives* (pp. 1–22). Westport: Praeger.

Lambert, M. C., Rowan, G. T., Lyubansky, M., & Russ, C. M. (2002). Do problems of clinic-referred African-American children overlap with the child behavior checklist? *Journal of Child and Family Studies, 11*(3), 271–285. doi:10.1023/A:1016816005732.

Lambert, M. C., Schmitt, N., Samms-Vaughan, M., An, J. S., Fairclough, M., & Nutter, C. A. (2003). Is it prudent to administer all items for each child behavior checklist cross-informant syndrome? Evaluating the psychometric properties of the youth self-report dimensions with confirmatory factor analysis and item response theory. *Psychological Assessment, 15*(4), 550–568. doi:10.1037/1040-3590.15.4.550.

Lambert, M. C., Rowan, G. T., Kim, S., Rowan, S. A., An, J. S., Kirsch, E. A., & Williams, O. (2005). Assessment of behavioral and emotional strengths in black children: Development of the behavioral assessment for children of African heritage. *Journal of Black Psychology, 31*(4), 321–351. doi:10.1177/0095798405278409.

Last, C. G., & Perrin, S. (1993). Anxiety disorders in African-American and white children. *Journal of Abnormal Child Psychology, 21*(2), 153–164. doi:10.1007/BF00911313.

Mann, E. M., Ikeda, Y., Mueller, C. W., & Takahashi, A. (1992). Cross-cultural differences in rating hyperactive-disruptive behaviors in children. *The American Journal of Psychiatry, 149*(11), 1539–1542.

McDermott, P. A., & Spencer, M. B. (1997). Racial and social class prevalence of psychopathology among school-age youth in the United States. *Youth & Society, 28*(4), 387–414. doi:10.1177/0044118X97028004001.

McDermott, P. A., Marston, N. C., & Stott, D. H. (1993). *Adjustment scales for children and adolescents*. Philadelphia: Edumetric and Clinical Science.

Meikamp, J. (2003). Review of the Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior

Inventory-revised. In *The fifteenth mental measurements yearbook*. http://www.unl.edu/buros/. Accessed 25 Mar 2013.

Miller, T. W., Nigg, J. T., & Miller, R. L. (2009). Attention deficit hyperactivity disorder in African American children: What can be concluded from the past ten years? *Clinical Psychology Review, 29*(1), 77–86. doi:10.1016/j.cpr.2008.10.001.

Neal, L. V. I., McCray, A. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The effects of African American movement styles on teachers' perceptions and reactions. *The Journal of Special Education, 37*(1), 49–57. doi:10.1177/00224669030370010501.

Neblett, E. W. Jr., White, R. L., Ford, K. R., Philip, C. L., Nguyên, H. X., & Sellers, R. M. (2008). Patterns of racial socialization and psychological adjustment: Can parental communications about race reduce the impact of racial discrimination? *Journal of Research on Adolescence, 18*(3), 477–515. doi:10.1111/j.1532-7795.2008.00568.x.

Nolan, E. E., Gadow, K. D., & Sprafkin, J. (2001). Teacher reports of DSM-IV ADHD, ODD, and CD symptoms in school children. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(2), 241–249. doi:10.1097/00004583-200102000-00020.

Pastor, P. N., & Reuben, C. A. (2005). Racial and ethnic differences in ADHD and LD in young school-age children: Parental reports in the national health interview survey. *Public Health Reports, 120,* 383–392.

Puig, M., Lambert, M. C., Rowan, G. T., Winfrey, T., Lyubansky, M., Hannah, S. D., & Hill, M. F. (1999). Behavioral and emotional problems among Jamaican and African American children, ages 6 to 11: Teacher reports versus direct observations. *Journal of Emotional and Behavioral Disorders, 7*(4), 240–250. doi:10.1177/106342669900700406.

Reid, R., DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Rogers-Adkinson, D., Noll, M., & Riccio, C. (1998). Assessing culturally different students for attention deficit hyperactivity disorder using behavior rating scales. *Journal of Abnormal Child Psychology, 26*(3), 187–198. doi:10.1023/A:1022620217886.

Reid, R., Casat, C. D., Norton, H. J., Anastopoulos, A. D., & Temple, E. P. (2001). Using behavior rating scales for ADHD across ethnic groups: The IOWA Conners. *Journal of Emotional and Behavioral Disorders, 9*(4), 210–218. doi:10.1177/106342660100900401.

Rescorla, L. A. (2009). Rating scale systems for assessing psychopathology: The Achenbach system of empirically based assessment (ASEBA) and the behavior assessment system for children-2 (BASC-2). In J. L. Matson, F. Andrasik, & M. L. Matson (Eds.), *Assessing childhood psychopathology and developmental disabilities* (pp. 117–149). New York: Springer. doi:10.1007/978-0-387-09528-8_5.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment system for children (BASC-2)* (2nd ed.). Circle Pines: AGS.

Roberts, R. E., Roberts, C. R., & Xing, Y. (2006). Prevalence of youth-reported DSM-IV psychiatric disorders among African, European, and Mexican American adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(11), 1329–1337. doi:10.1097/01.chi.0000235076.25038.81.

Ruchkin, V., Jones, S., Vermeiren, R., & Schwab-Stone, M. (2008). The strengths and difficulties questionnaire: The self-report version in American urban and suburban youth. *Psychological Assessment, 20*(2), 175–182. doi: 10.1037/1040-3590.20.2.175.

Rutter, M. (1967). A Children's Behaviour Questionnaire for completion by teachers: Preliminary findings. *Journal of Child Psychology and Psychiatry, 8*(1), 1–11. doi:10.1111/j.1469-7610.1967.tb02175.x.

Seaton, E. K., Caldwell, C. H., Sellers, R. M., & Jackson, J. S. (2008). The prevalence of perceived discrimination among African American and Caribbean Black youth. *Developmental Psychology, 44*(5), 1288–1297. doi:10.1037/a0012747.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. (2000). NIMH diagnostic interview schedule for children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(1), 28–38. doi:10.1097/00004583-200001000-00014.

Smitherman, G. (2004). Talkin and testifyin: Black English and the black experience. In. R. L. Jones (Ed.). *Black psychology* (4th ed., pp. 295–311). Hampton: Cobb & Henry.

Stein, S. (2007). Review of the behavior assessment system for children, Second Edition. In *The seventeenth mental measurements yearbook*. http://www.unl.edu/buros/. Accessed 25 Mar 2013.

U.S. Census Bureau. (2010). *Hispanic population: 2010*. http://www.census.gov/prod/cen2010/briefs/c2010br-06.pdf. Accessed 25 Mar 2013.

Williams, R. L. (1972). *The BITCH-100: A cultural specific test*. St Louis: Williams

Worrell, F. C. (2009). Psychological health in school-age populations. In H. A. Neville, B. M. Tynes, & S. O. Utsey (Eds.), *Handbook of African American psychology* (pp. 351–362). Thousand Oaks: Sage Publications.

Youngstrom, E. (2008). Evidence-based strategies for the assessment of developmental psychopathology: Measuring prediction, prescription, and process. In W. E. Craighead, D. J. Miklowitz, & L. W. Craighead (Eds.), *Psychopathology: History, diagnosis, and empirical foundation* (pp. 34–77). Hoboken: John.

Zamarippa, M. X., & Lerma, E. (2013). School-based assessment with Latina/o children and adolescents. In L. T. Benuto (Ed.), *Guide to psychological assessment with Hispanics* (pp. 335–349). New York: Springer.

Zimmerman, R. S., Khoury, E. L., Vega, W. A., Gil, A. G., & Warheit, G. J. (1995). Teacher and parent perceptions of behavior problems among a sample of African American, Hispanic, and non-Hispanic white students. *American Journal of Community Psychol*ogy, 23(2), 181–197. doi:10.1007/BF02506935.

9

Assessment of Generalized Anxiety Disorder, Panic, and Phobias with African Americans

L. Kevin Chapman, Ryan C. T. DeLapp and Allyn Richards

Generalized anxiety disorder (GAD), panic disorder with and without agoraphobia (PDA/PDWA), specific phobia, and social phobia are anxiety disorders that are commonly experienced among African Americans (Himle et al. 2009). Although burgeoning, the current empirical literature remains relatively sparse as it relates to culturally sensitive assessment of anxiety disorders. In some cases, cross-cultural comparisons denote differences between African Americans and non-Hispanic Whites (e.g., prevalence rates, Himle et al. 2009; Paradis et al. 1993; mean differences on measures of anxiety, Scott et al. 2002). However, a potential confusion in these findings is whether the psychometric properties of these assessment tools are valid in ethnically heterogeneous samples (see Melka et al. 2010) since the majority of "gold standard" instruments have been validated in non-Hispanic Whites. As such, there is a plethora of cultural factors (e.g., perceived stress, cultural mistrust, racial identity, culturally endemic explanatory models) that should be considered in order to improve the clinical utility of anxiety measures for use with African Africans. Along these lines, the goal of this chapter is to provide an overview of cultural considerations in the assessment of GAD, panic disorder (PD), and the phobias and summarize extant literature delineating culturally specific findings on "gold standard" assessments of the aforementioned anxiety disorders. Detailed information regarding all relevant measures is provided below organized by disorder and then by assessment type (e.g., interview rating scales, self-report measures). Table 9.1 provides a summary of recommendations for the specific measures made throughout the chapter.

Generalized Anxiety Disorder

Generalized anxiety disorder (GAD) encompasses a tendency to excessively worry about future, negative events, and/or general life concerns (e.g., finances, family matters) in a fashion that causes considerable distress and a perceived inability to control one's thoughts (Barlow 2004). Associated elements of GAD include sleep disturbances, restlessness, fatigue, difficulty in concentrating, and muscle tension (American Psychiatric Association 2000). Evidence supports that individuals with debilitating worry maintain an attentional bias for threatening information in their environment, whereas the alternative is not supported in those who worry less (Carter et al. 2005). More recently, Watson (2005) grouped GAD and depression under "distress-based disorders" to account for their overlap in maintaining negative

L. K. Chapman (✉) · R. C. T. DeLapp · A. Richards
Department of Psychological and Brain Sciences,
University of Louisville, 2301 South 3rd Street,
Louisville, KY 40292, USA
e-mail: kevin.chapman@louisville.edu

R. C. T. DeLapp
e-mail: rcdela01@exchange.louisville.edu

A. Richards
e-mail: aerich04@louisville.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_9, © Springer Science+Business Media New York 2015

Table 9.1 Recommendations for using assessments with African Americans

Assessment name	Disorder assessed	Recommendations
Anxiety Diagnostic Interview Schedule-IV	GAD	*Strength*: Controls for underreporting cognitive symptoms *Limitation*: Role of cultural mistrust (Hunter and Schmidt 2010)
Hamilton Rating Scale	GAD	*Strength*: Controls for underreporting cognitive symptoms *Limitation*: Role of cultural mistrust
Penn State Worry Questionnaire	GAD	*Strength*: Identifies the excessiveness of worry and best discriminates worry in clinical populations (Hambrick et al. 2010) *Limitations*: Not an assessment of content of worry and poorly discriminates worry in subclinical populations (Hambrick et al. 2010)
Worry Domains Questionnaire	GAD	*Strength*: Describes content of worry *Limitations*: Finite number of life events examined, which may overlook culturally specific worries in African Americans
General Anxiety Disorder-Questionnaire-IV	GAD	Strengths: Examines worry according to DSM-IV criterion Incremental and divergent validity across ethnic groups (Robinson et al. 2010)
Anxiety Diagnostic Interview Schedule-Version IV	Social phobia	*Strength*: Controls for underreporting cognitive symptoms Assesses symptoms of social phobia according to DSM-IV criteria *Limitation*: Role of cultural mistrust (Hunter and Schmidt 2010) Has not been exclusively examined in African American samples
Liebowitz Social Anxiety Scale	Social phobia	*Strengths*: Good internal consistency and temporal stability in community-based African American sample (Beard et al. 2011) Can be used to diagnose social phobia *Limitations*: Further exploratory factor analysis needed to establish its utility in African American samples (Beard et al. 2011) Role of cultural mistrust (Hunter and Schmidt 2010)
Social Interaction Anxiety Scale and Social Phobia Scale	Social phobia	*Strengths*: Good internal consistency in a community-based African American sample (Chapman et al., in review) *Limitations*: Extant research has found that the SIAS did not explain the variance in social anxiety among their African American sample (Hambrick et al. 2010)
Fear of Negative Evaluation Scale and Social Avoidance Distress Scale	Social phobia	*Strengths*: Relatively easy to administer Short (28 items) *Limitations*: Factor variance
Positive Affect and Negative Affect Scale	Social phobia	Strengths: Screening tool for anxiety and social phobia diagnoses in African American women. Limitations: Cannot be used solely to diagnose social phobia Does not convey severity or content of social anxiety
Social Phobia and Anxiety Inventory	Social phobia	Strengths: Differentiates between social anxiety and agoraphobia Limitations: SPAI and SPAI-23 have not been exclusively examined in African American samples
Anxiety Diagnostic Interview Schedule-IV	Panic disorder	Strength: Controls for underreporting cognitive symptoms *Limitation*: Role of cultural mistrust (Hunter and Schmidt 2010) Psychometric properties of panic section not examined exclusively in African American sample

Table 9.1 (continued)

Assessment name	Disorder assessed	Recommendations
Panic Disorder Severity Scale	Panic disorder	*Strength*: Has demonstrated convergent validity with both clinical interviews and self-report *Limitation*: Has yet to be evaluated in African American samples Due to a tendency to report somatic symptoms, African Americans may be rated inaccurately high on the scales that assess physical sensations and complaints
Agoraphobic Cognitions Questionnaire	Panic disorder/ agoraphobia	*Strength*: Assesses cognitive symptoms of panic and agoraphobia that are traditionally underreported by African American clients *Limitation*: Psychometric properties and clinical utility have not been examined in African American samples
Anxiety Sensitivity Index	Panic disorder	*Strength*: The four-factor model found by Arnau et al. (2009) was found to have strong reliability and validity in comparative sample of African American and non-Hispanic White college students *Limitation*: Anxiety Sensitivity Index-Revised (Taylor and Cox 1998) may have greater predictive validity Psychometrics have yet to be examined exclusively in African American and clinical samples
Brief Panic Disorder Screen	Panic disorder	*Limitations*: Demonstrates significantly weaker psychometric properties in African American samples compared to non-Hispanic white samples BPDS may not be appropriate in the assessment and determination of panic disorder and panic disorder with agoraphobia in African Americans
Albany Panic and Phobia Questionnaire	Panic disorder/ agoraphobia	*Strength*: May be an internally consistent measure of agoraphobia and interoceptive fears related to panic disorder in African Americans *Limitations*: Validity and clinical utility (e.g., cutoff scores) in assessment of panic disorder need to be determined in African Americans
Screen for Child Anxiety Related Emotional Disorders-Panic Subscale	Panic disorder	*Strength*: Has demonstrated sound psychometric properties in African American child samples The total subscale appears to have clinical utility in the prediction of clinical anxiety *Limitations*: The psychometric properties and clinical utility of the SCARED need to be determined specifically within the context of panic disorder in African American children
Child Anxiety Sensitivity Index	Panic disorder	*Strength*: Research has supported relationship between anxiety sensitivity and panic in African American youth Demonstrates sound psychometric properties in African American child samples Evaluates cognitive symptoms associated with panic that may be traditionally underreported in African Americans *Limitation*: Discrepant factor structures between African American and non-Hispanic White samples may indicate that the CASI does not fully capture construct of anxiety sensitivity within this population
Anxiety Diagnostic Interview Schedule-IV	Specific phobia	*Strength*: Assesses full range of symptoms of specific phobia according to DSM-IV criteria Assesses cognitive symptoms that may be underreported Has adult, parent, and child versions Limitation: Psychometric properties and clinical utility specific to specific phobia diagnosis in African Americans have yet to be examined

Table 9.1 (continued)

Assessment name	Disorder assessed	Recommendations
Fear Survey Schedule-Second Edition	Specific phobia	*Strength*: One of the few self-report measures that assesses multiple domains of fear related to specific phobia Factor structure appears to be similar between both college and community samples of African American adults *Limitation*: Psychometric properties and clinical utility within the context of specific phobia in African Americans needs to be examined
Revised Fear Survey Schedule for Children	Specific phobia	*Strength*: One of the few self-report measures that assesses broad domains of fears related to specific phobia| Has demonstrated strong psychometric properties in an African American child sample *Limitations*: The predictive validity and clinical utility in the context of specific phobia diagnoses in African American children remains unknown

mood states and general distress that pervades across life experiences.

Though extant literature suggests that African Africans (1.37 %) have a significantly lower prevalence rate of GAD than non-Hispanic Whites (4.55 %; Himle et al. 2009), there are specific cultural considerations that could explain these findings. First, examiners should remain aware of the stigma associated with mental health in the African American community (Hunter and Schmidt 2010). As such, African Americans may be reluctant to report worry symptoms due to the psychological, verbally linguistic nature of this coping attempt. Moreover, we will describe the "gold standard" assessment options for examining GAD and delineate cultural considerations when administering each assessment measure with African Americans.

GAD Assessments

Interviewer Rating Scales

Anxiety Diagnostic Interview Schedule-Version IV (ADIS-IV-Client Version; Brown et al. 1994) The ADIS-IV-Client Version is a semistructured interview for adults that examines the presence and severity of anxiety and other mental illnesses according to the criteria outlined in the *Diagnostic and Statistical Manual of Mental Disorder-4th edition* (DSM-IV; American Psychiatric Association 1994). In the assessment of GAD, the ADIS-IV-Client Version measures the excessiveness and uncontrollability of worry across several domains (e.g., Minor Matters, Work/School, Family, Finances). Additional ratings include an assessment of the presence and severity of associated symptoms (e.g., restlessness, difficulty concentrating), the daily interference caused by worry, and the level of distress experienced due to the worries. Examiners provide a clinical severity rating (CSR), ranging from 0—"none" to 8—"severely disturbing/disabling," where a score of 4 or greater indicates meeting DSM-IV diagnostic criteria.

To date, no study has exclusively examined the psychometric properties of the GAD section of the ADIS-IV in African Americans. However, existing literature suggests that the ADIS-IV has excellent interrater reliability ($k = 0.85$ and 0.90) across all anxiety disorders in an exclusively African American adult sample (Chapman et al. 2012). Overall, an inherent strength of the ADIS-IV appears to be its close adherence to the guidelines of the DSM-IV criterion while allowing the clinician to further assess symptoms of GAD that may not be captured on self-report measures. However, examiners must be mindful of the influences of racial identity, ethnicity, acculturation, and their collective impact on the endorsement of worry symptoms in African American adults (for review, see Carter et al. 1996).

Finally, the ADIS-IV-Parent and Child (Silverman and Albano 1996) are the child versions of this diagnostic interview that include additional examinations of childhood-based disorders (e.g., attention-deficit/hyperactivity disorder, separation anxiety disorder) as well as more age-appropriate language. According to our review of the literature, these versions of the ADIS-IV have not been exclusively utilized to examine childhood GAD in African Americans.

Hamilton Anxiety Rating Scale (HARS; Hamilton 1959) The HARS is a 14-item scale that is designed to evaluate the severity of anxiety symptoms rather than the presence of a particular disorder (Barlow 2004); however, it is most commonly used in the assessment of GAD (Shear et al. 2001a). The HARS contains two subscales: psychic, which examines symptoms of apprehension and irritability, and somatic symptoms, which assesses autonomic arousal associated with anxiety. Previous literature supports the convergent validity of the HARS, but there is skepticism in its discriminant validity and interrater reliability (Barlow 2004). Subsequently, structural interview guides (SIG) have been created to improve the psychometric qualities of the HARS by standardizing its administration and interpretation (Bruss et al. 1994; Shear et al. 2001a). The most recently developed SIG by Shear and colleagues demonstrated slightly better interrater reliability compared to non-SIG HARS assessments within a predominantly non-Hispanic White clinical sample. According to our review of the extant literature, no study to date has examined the psychometric properties of the HARS for the assessment of worry in African Americans; however, the aforementioned considerations related to the ADIS-IV are suggested for the HARS.

Self-Report Scales

Penn State Worry Questionnaire (PSWQ; Meyer et al. 1990) The PSWQ is a 16-item measurement that assesses chronic worry on a 5-point Likert scale (1: "Not at all typical of me" and 5:"Very typical of me"). The purpose behind the PSWQ is to assess the excessiveness and intensity of worry rather than the specific content of or physical symptoms associated with the respondents worries (Barlow 2004). Based upon research supporting the temporal stability of scores as well as the general phrasing of the items (e.g., "I worry all the time"), PSWQ is considered an assessment of the emotional trait of worry instead of a momentary state of worry (Carter et al. 2005). In terms of its psychometric properties in African Americans, Chapman, Kertz, and Woodruff-Borden (2009b) found good internal consistency (a = 0.73) when using the PSWQ in an African American college sample as well as lower mean scores in comparison to non-White Hispanic students.

Notably, there have been discrepancies in the factor patterns between these groups. In particular, Carter et al. (2005) examined the factor structure of the PSWQ across ethnic groups and found that African American students were uniquely described by a three-factor structure that contained items describing General Worry (e.g., "I am always worrying about something"), Worry Absence (e.g., "I never worry about anything"), and Worry Dismissal (e.g., "If I don't have enough time to do everything, I don't worry") whereas non-Hispanic Whites only contained a two-factor solution, General Worry and Worry Absence. The main distinction between these factor structures was that the item, "I never worry about anything," loaded on the Worry Absence factor for non-Hispanic Whites and on the Worry Dismissal factor for African Americans. Despite describing discrepant factor patterns between ethnic groups, researchers propose that the common General Worry factor was a sound assessment of worry across both samples because: (1) the factor contains 11 items directly focused on the presence of excessive worry which is the hallmark feature of GAD (see Table 9.2) and (2) the factor demonstrated good convergent validity with the State Trait Anxiety Inventory (STAI; Spielberger et al. 1983), indicating the General Worry factor and the STAI are measuring different facets of the same higher-order construct (i.e., anxiety). Contrarily, the General Worry factor exhibited less than favorable discriminant validity with the Beck Depression Inventory (BDI; Beck and Steer 1987) in African American students, which highlights the overlap between the symptoms of these disorders and suggests that an additional assessment may be needed to distinguish

Table 9.2 Eleven PSWQ items comprising the "General Worry" factor that similarly described worry across samples of African American and non-Hispanic young adults. (Carter et al. 2005)

7.	I am always worrying about something
13.	I notice that I have been worrying about things
15.	I worry all the time
5.	I know I should not worry about things, but I just cannot help it
6.	When I am under pressure I worry a lot
14.	Once I start worrying, I cannot stop
4.	Many situations make me worry
9.	As soon as I finish one task, I start to worry about everything else I have to do
12.	I have been a worrier all my life
2.	My worries overwhelm me
16.	I worry about project until they are done

GAD and depression in African Americans when using the PWSQ.

Furthermore, more recent literature proposes three additional cultural considerations when administering the PSWQ to African Americans. First, Chapman et al. (2009) found that psychological distress and perceived control predicted PSWQ scores in both samples; however there were significant differences in terms of which construct contributed most within each pathway. For African Americans, psychological distress was more characteristic of their worry than perceived control, whereas low perceived control contributed more to worry for non-Hispanic Whites. Though it is unclear whether these findings are generalizable to clinical samples, this study suggests that psychological distress is an especially salient component of worry for African Americans that should be considered when assessing GAD. Also, Rucker, West, and Roemer (2010) found that the relationship between perceived racism and PSWQ scores was accounted for by intolerance of uncertainty (or the tendency to negatively interpret information within ambiguous situations in way that leads to anxiety-provoking cognitive, emotional, and behavioral states). This finding suggests that a useful adjunct to worry assessment in African Americans is to further examine cognitions of uncontrollability and unpredictability of future events, which are endemic to both anxiety and intolerance of uncertainty. Third, Hambrick et al. (2010) concluded that PSWQ is most effective in discriminating worry within clinical populations of African Americans and cautioned its use in samples with subclinical worry after comparing the differential item functioning of the PSWQ across African American, non-Hispanic White, and Asian American college students.

Finally, the PSWQ has a modified version that can be used in children and adolescents. The Penn State Worry Questionnaire-Children (PSWQ-C; Chorpita et al. 1997) is a 14-item self-report assessment of the tendency to worry. The respondents are prompted to rate how often they endorse certain worries on 4-point Likert scale (i.e., never, sometimes, often, and always). According our review of literature, no studies have established cultural considerations exclusively for African American youth.

Worry Domains Questionnaire (WDQ; Tallis et al. 1992) The WDQ is a 25-item assessment of the content and severity of worry across five subscales: Relationships (e.g., "that I will lose close friends"), Lack of Confidence (e.g., "that I feel insecure"), Aimless Future (e.g., "that I'll never achieve my ambitions"), Work Incompetence (e.g., "that I make mistakes at work"), and Financial (e.g., "that my money will run out"). For a comprehensive assessment of worry, Barlow (2004) suggests that the WDQ be administered in conjunction with the PSWQ to both examine the content and excessiveness of the respondent's worry. Evidence supports the WDQ's temporal stability of scores across 4 weeks, its convergent validity with peer reported worry, and its internal consistency across all five subscales (Stober 1998). In terms of cultural considerations when using the WDQ, one study compared the excessiveness (i.e., measured by PSWQ) and content (i.e., measured by WDQ) of worry across African American, non-Hispanic White, and Asian American college students (Scott et al. 2002). Although ethnic groups did not exhibit any mean differences on the PSWQ, African Americans endorsed significantly lower scores across each content subscale (except for the Financial domain) as compared to their counterparts. In particular, African Americans reported the greatest worry in the Financial and Relationship domains whereas their lowest scores fell in the content area of Aimless Future. Researchers conclude that their findings suggest that African Americans worry less across WDQ subscales (except for financial worries), however an alternative explanation lies in whether the WDQ comprehensively examines the worry domains in African Americans. In other words, it is important to note that the WDQ, a measure that only assesses a finite number of life events, may not entirely reflect the spectrum of worries that African American respondents experience, which could account for the lower scores endorsed by this population. Given this possibility, examiners should consider that African Americans may experience excessive and/or uncontrollable worry in other culturally specific areas not captured by the five subscales of the WDQ. Additionally, Scott and colleagues found good internal consistency across each ethnic group, yet they propose that further empirical study is needed to replicate their findings in clinical samples as well as examine the validity of the WDQ cross-culturally.

Generalized Anxiety Disorder Questionnaire-IV (GAD-Q-IV; Newman et al. 2002) The GAD-Q-IV is specifically designed, unlike the PSWQ and WDQ, to examine GAD according to criteria outlined in the DSM-IV (APA 2000). Specifically, the GAD-Q-IV is a 9-item self-report measure that assesses the expression and interference of excessive worry as well as somatic symptoms associated with worry. Past literature has examined the use of two GAD-Q-IV scoring systems: (1) criterion matching, or comparing the endorsed items to DSM-IV criterion to diagnose GAD, and (2) dimensional scoring, or using a cutoff score of approximately 6 to diagnosis GAD. Luterek et al. (2002) found that both scoring methods exhibited good specificity (>90%) in their community-based sample (17% African American), but the criterion matching method had slightly lower sensitivity (77.4%) than the dimensional method (>90%). Notably, due to their control group mainly consisting of individuals without an Axis I disorder (51 out of 53), researchers still question the specificity of the dimensional method. Aside from its clinical utility, the GAD-Q-IV has shown good reliability with structured interviews (Barlow 2004; Robinson et al. 2010). Regarding the cultural considerations for using the GAD-Q-IV, Robinson and colleagues compared its psychometric characteristics across ethnically diverse college and clinical samples. Results indicated no mean differences in addition to a consistent one-factor structure across African American, non-Hispanic White, Hispanic American, and Asian American groups. Furthermore, the GAD-Q-IV demonstrated good incremental validity by predicting more variance in PSWQ scores compared to general anxiety (i.e., BAI) and depression (i.e., BDI) measures. Additionally, divergent validity was established in both college and clinical samples as evidenced by the GAD-Q-IV maintaining a non-significant relationship with the Panic Disorder Severity Scale (PDSS; Shear et al. 1997). Based upon their findings, researchers concluded that the GAD-Q-IV "is measuring the diagnostic construct of GAD uniformly across groups." (Robinson et al. 2010, p. 258)

Social Phobia

Social anxiety is the distress experienced in social interactions where there is a fear of being negatively evaluated or that you will behave in a fashion that may cause humiliation or embarrassment. This form of anxiety is often coupled with heightened somatic arousal (e.g., increase heart rate, slowed breathing, nausea), and may become debilitating when one develops avoidance behaviors (e.g., repeatedly refusing to speak in front of groups) or one endures social situations with considerable distress. In identifying pathological social fear (i.e., social phobia), it is important to consider the degree of avoidance and distress that characterizes one's social experiences as well as how their social anxiety is impacting their daily functioning (e.g., social, occupational, academic). Extant literature has yielded mixed results when comparing the prevalence of social anxiety cross-culturally. Whereas Brown and Eaton (1986) found higher prevalence of social phobia in a community sample of African Americans (5.6%) as compared to a non-Hispanic White sample (2.6%), the National Survey of American Life and National Comorbidity Survey-Replication reported less instances of social phobia (Himle et al. 2009; Ruscio et al. 2008). Furthermore, Himle and colleagues found that the greatest risk factor for developing social phobia in African Americans was being younger than 20 years of age.

Several social anxiety assessments have been validated in academic settings and therefore their generalizability to clinical and/or community-based African American populations must be considered. However, there is evidence supporting the relevance of examining varied levels of social anxiety within school settings given the highly performance-based and rigorous evaluative dynamics in higher education. In particular, stereotype threat, a phenomenon describing awareness to an identity (e.g., race, gender) in a situation where that identity is stereotyped to perform poorly, is especially pertinent to social anxiety experiences of African Americans. Although there has not been a clear empirical explanation for the relationship between stereotype threat and social anxiety, nascent research suggests that an awareness of negative stereotypes concerning one's ethnic group is predictive of poorer performance on tests in African American students (Steele and Aronson 1995). Conceptually, the effects of stereotype threat could be generalized to other performance-based situations outside the realm of testing, where the consciousness of race-based stereotypes (e.g., African Americans are not highly educated beings) could explain certain elements of significant distress and/or situational avoidance (e.g., not wanting to speak up during meetings in front of predominantly White colleagues) in African Americans. Furthermore, perceived racism, or the suspicion that one has been discriminated against, can also uniquely impact the social anxiety of African Americans, by acting as a precipitating and/or reinforcing factor in the manifestation of socially based fears.

In this section, the "gold standard" options for examining social phobia will be discussed in addition to empirically supported cultural considerations when administering each assessment option in African Americans.

Interviewer-Rating Scales

ADIS-IV As previously discussed, the ADIS-IV (Brown et al. 1994) is a commonly used diagnostic interview that allows for the differentiation of anxiety and related disorders according to the criterion set by the DSM-IV. The social phobia section of the ADIS-IV includes an array of social settings (e.g., parties, speaking with unfamiliar people) that prompts respondents to rate their degree of fear and avoidance of the endorsed items (for additional considerations, see section on ADIS-IV and GAD). Although the psychometric qualities of this section of the ADIS-IV has yet to be examined in African Americans, the ADIS-IV has demonstrated excellent interrater reliability in exclusively African American adult samples (Chapman et al. under review; Petrie 2013).

As aforementioned, given that the social phobia section of the ADIS-IV closely aligns with DSM-IV criteria, diagnoses made using this assessment should accurately reflect the presence and severity of respondent's social fears. However, a limitation of the assessment tool is that it does directly prompt for culturally based experiences with discrimination, and therefore requires the examiner to separately assess the role that these experiences may have on the respondent's social fears.

Liebowitz Social Anxiety Scale (LSAS; Liebowitz 1987) The LSAS contains 24 items that assess the severity of anxious symptoms and avoidant behaviors across various social situations. Respondents rate the severity of their fear and avoidance on a 4-point Likert scale across two subscales, social-interactional situations (i.e., 13 items) and performance situations (i.e., 11 items). There is no existing literature that describes the use of the LSAS in child populations; however it has demonstrated good psychometric properties and clinical utility in adults (Heimberg et al. 1999). Notably, a recent study utilizing the LSAS in a community sample of African Americans found good internal consistency in total and original subscale scores as well as the temporal stability of total and subscale scores over a year (Beard et al. 2011). However, researchers found high intercorrelations between the fear and avoidance rating scales, which yielded redundant information in the African American sample. Beard et al. (2011) further examined the original and extant factor structures of the LSAS (i.e., original two subscales—Liebowitz 1987; 4 subscales—Safren et al. 1999; 5 subscales—Baker et al. 2002) and concluded that three of the four-factor structures (including the original) resulted in poor model fit, although the Safren (1999) model resulted in minimally acceptable fit. Taken together, these results suggest that further exploratory factor analyses are warranted in order to underscore the utility of the LSAS in African American samples (see Beard et al. 2011).

Self-Report Scales

Social Interaction Anxiety Scale (SIAS) and Social Phobia Scale (SPS; Mattick and Clark 1998) Both the SIAS and the SPS are 20-item assessments that utilize a 4-point Likert scale (i.e., 0—"Not at all" to 4—"Extremely") to examine anxiety-related reactions across a number of social situations (i.e., SIAS) and fears of being scrutinized by others (i.e., SPS). Evidence supports that high scores on the SIAS are related to more severe social anxiety across various social situations, whereas higher scores on the SPS are associated with a heightened focus on symptoms of somatic arousal as well as the potential consequences of these symptoms (Barlow 2004). Both measures have demonstrated good reliability and validity in past studies (Barlow 2004).

Specifically, two studies to date have examined the psychometric qualities of the SIAS and SPS exclusively in African American samples (Hambrick et al. 2010; Chapman et al. in review). Hambrick and colleagues examined the mean differences and differential item responses on the SIAS across African American, Asian American, and non-Hispanic White subjects and found that African Americans had significantly lower total scores amongst the entire sample, with the largest discrepancy between African Americans (20.16) and Asian Americans (27.01). Also, researchers determined that item responses across the ethnic groups were disparate and, most notably, that the responses on the SIAS did not effectively differentiate African American participants with varying degrees of social interaction anxiety (Hambrick et al. 2010). Ultimately, researchers concluded that using the SIAS to compare social interaction anxiety across diverse ethnic groups must be done with caution based upon its inability to explain the variance in their African American sample. Despite these limitations, Chapman et al. (under review) found good internal consistency (SIAS a = 0.92; SPS a = 0.94) as well as established cut scores for identifying clinically significant social fear using the SIAS and SPS in an exclusively African American sample. The socially anxious subjects in the sample had an average SIAS score of 32.26 and SPS score of 20.39. Notably, researchers found that a SIAS score of 15 achieved 85 % sensitivity (e.g., true positives) and 82 % specificity (e.g., true negatives) within the current sample. Furthermore, an SPS of 6 attained 74 % sensitivity and 77 % of specificity.

Fear of Negative Evaluation Scale (FNES) and Social Avoidance Distress Scale (SADS) The *FNES and the SADS* were developed concurrently to assess different aspects of social phobia (Watson and Friend 1969) The FNES is a 30-item self-report assessment that examines concern about social or public scrutiny based upon "True" or "False" responses. Sample questions include "I feel very upset when I commit some social error" and "I often worry that I will say or do the wrong things." The SADS is a 28-item self-report instrument that measures the degree of discomfort in social situations and likelihood to avoid social interaction based upon "True" or "False" responses. Sample items include "I feel relaxed even in unfamiliar social situations" and "I often feel nervous or tense in casual get-togethers in which both sexes are present."

In terms of psychometric properties of these measures in African Americans, a recent investigation identified discrepant factor structures between African Americans and non-Hispanic Whites (Melka et al. 2010). Researchers found that several items on the FNES, including "I am unconcerned even if I know people are forming an unfavorable impression of me," "the opinions that important people have of me cause me little concern," "I react very little when other people disapprove of me," "I am usually confident that others will have a favorable impression of me," and "I am unconcerned even if I know people are forming an unfavorable impression of me," were not salient concerns in their African American college sample. On the SADS, the items that did not appear pertinent to African Americans were "I often find social occasions upsetting" and "I am seldom at ease in a large group of people." Although there are concerns about the generalizability of these findings to a more demographically diverse sample (e.g., community-based or clinical populations), Melka et al. (2010) concluded that items reflecting indifference to negative evaluation and degrees of comfort in large social interactions were not representative of the social anxiety in their African American sample. Most importantly, these findings reiterate the importance of utilizing culturally sensitive assessment items (Melka et al. 2010) because when the average scores were reanalyzed using the refined SAD and FNE items, African Americans were no longer significantly higher than the other sample on the SAD (but still remained lower on the FNE). Ultimately, by including items that are most descriptive of social concerns about negative evaluations (i.e., FNE) and apprehensions about social interactions (i.e., SAD) among African Americans, it improves the precision of assessing social anxiety across ethnicities.

Positive Affect and Negative Affect Scale (PANAS; Watson et al. 1988) The PANAS includes 20 items measuring dimensions of positive (i.e., 10 items) and negative affect (i.e., 10 items). Specifically, negative affectivity describes distressful mood states that are usually accompanied by anger, sadness, guilt, or disgust, whereas positive affectivity defines emotional states of happiness, high energy, and satisfaction. Both of these concepts operate on a continuum where low negative affect characterizes calmness and serenity, and low positive affect represents sadness and lethargy (Watson et al. 1988). The PANAS employs a 5-point Likert scale, ranging from "very slightly or not at all" to "severely," to examine the respondent's level of affectivity. This assessment tool has been used to examine affectivity in the moment, over the span of a day, a few days, a week, a few weeks, a year, and generally (Watson et al. 1988). The PANAS has been shown to be reliable and valid in various samples, including non-Hispanic White college students (Watson et al. 1988).

Evidence supports an association between social anxiety and low positive affectivity that is similar to depression yet uncharacteristic of other anxiety disorders. In particular, high levels of social anxiety have been found to be associated with lower positive affect on the PANAS compared to individuals with lower levels of social anxiety (Vittengl and Holt 1998). Though this relationship has been substantiated in non-Hispanic Whites, one study examined the clinical utility of the PANAS in differentiating community-based African American female adults diagnosed with and without an anxiety disorder as well as its ability to predict those who had social phobia diagnoses. In terms of identifying an anxiety diagnosis within the sample, researchers determined that the PANAS established cutoff scores for both scales in predicting overall anxiety diagnosis (>11 on negative scale; <35 on positive scale) and social phobia diagnosis (>13 on negative scale; <34 on positive scale). Petrie et al. (2013) demonstrate that the PANAS can be used as a screening tool to establish when a further assessment of pathological anxiety and, and more specifically, social phobia is warranted in African American females. However, because this assessment tool does not directly examine the content or severity of social fears, a supplemental measure of social phobia that more closely adheres to DSM-IV criteria should be included before arriving to a social phobia diagnosis. In terms of assessing children and adolescents, a 27-item self-report assessment called the PANAS-C (Laurent et al. 1999) was developed to examine affectivity in youth. Similar to the adult version, this child measure uses a 5-point Likert scale that prompts respondents to indicate how they have felt within last 2-week period. Despite a small sampling of African American youth, extant literature supports the link between positive affectivity (PA) and social phobia in youth as evidenced by the PA subscale of the PANAS-C significantly predicting social anxiety scores (Hughes and Kendall 2009). However, research is needed to establish cultural considerations when assessing social phobia in African American youth using the PANAS-C.

Social Phobia and Anxiety Inventory (SPAI; Turner et al. 1989) The SPAI is a 109-item assessment designed to examine physical, cognitive, and behavioral manifestations of distress across a number of anxiety-provoking situations (e.g., reading aloud, attending social gatherings) on a 7-point Likert scale. Uniquely, the SPAI includes an agoraphobia subscale, which allows for the differentiation between socially specific avoidant and generally situational avoidant respondents. However, literature has not remained consistent on the best way to utilize these subscales to accurately assess the severity of the respondent's social phobia. Suggestions include using the difference score between subscales (i.e., social phobia versus agoraphobia), solely using the social phobia subscale, or using either of the aforementioned options depending on the examiner's goal for the assessment (e.g., treatment outcome measure, general assessment of social phobia, or to distinguish symptoms across groups; for further review, see Barlow 2004). Evidence supports the psychometric quality of the SPAI as demonstrated by its greater diagnostic precision (e.g., sensitivity, specificity, and discriminability) compared to the SIAS and SPS (as cited in Schry et al. 2012). Recently, the development of an abbreviated version of the SPAI, the Social Phobia and Anxiety Inventory-23 (SPAI-23: Roberson-Nay et al. 2007), reduced the time burden placed on respondents while still maintaining a 2-factor structure despite only containing 23 of the original items. The psychometric quality of the SPAI-23 has been established by discriminating social phobia from other anxiety disorders and scores are shown to be highly correlated with the full SPAI in college students (as cited by Schry et al. 2012). For further psychometric information regarding the SPAI-23, see Schry et al. 2012.

According to our review of literature, no studies have exclusively investigated the SPAI or SPAI-23 in African Americans; however several studies have utilized the Social Phobia and Anxiety Inventory-Children (SPAI-C—Beidel et al. 1995) to assess social anxiety in African American youth. The SPAI-C is a 26-item assessment that maintains a similar approach to assessing social phobia in children aged 8–17. Evidence has yielded mixed findings as to how social phobia compares between African American and non-Hispanic White youth (Beidel et al. 2000; Ferrell et al. 2004; McLaughlin et al. 2007). Notably, among these studies, McLaughlin et al. (2007) included the largest sampling of African Americans ($n=141$) and found no differences between African Americans, non-Hispanic Whites, Hispanic Americans when considering overall total mean and scores across gender. For a more comprehensive summary of available child measures, consider the review article by Tulbure, Szentagotai, Dobrean, and David (2012). Overall, the bridge between the child and adult versions of SPAI are difficult to interpret despite both assessments including a similar operationalization of social phobia. In terms of the assessment of African American adults using the SPAI or SPAI-23, further research is needed to identify any cultural considerations when administering these tools.

Panic Disorder

Panic disorder (PD) is characterized by the experience of recurrent, unexpected panic attacks. These attacks are defined as discrete periods of intense fear in which at least four symptoms (out of 13) develop abruptly and climax within 10 min (American Psychiatric Association 2000). The symptoms of panic attack include physiological arousal, such as sweating, shortness of breath, trembling, and heart palpitations, as well as cognitive experiences, such as derealization, fear of losing control, and fear of "going crazy." In addition to experiencing several of these panic attack symptoms, an individual with PD must demonstrate a persistent fear regarding the potential for future attacks. Research has indicated that African Americans experience similar to lower lifetime prevalence rates of PD (i.e., 1.2–3.9 %) as compared to their non-Hispanic White counterparts (Breslau et al. 2006; Horwath et al. 1994). Previous work suggests that PD may be underdiagnosed in African Americans (Paradis et al. 1992).

One potential reason for these mixed findings may be the differential patterns of manifestation of PD in African Americans as compared to non-Hispanic Whites that have been consistently supported by the literature. Specifically, research has demonstrated that African Americans may be more likely to endorse the somatic symptoms of PD, such as tingling hands and feet and hot and cold flashes, and more overall symptoms (Horwath et al. 1994). Nascent studies have suggested that African Americans with a clinical diagnosis of PD may experience specific symptoms (i.e., tingling and numbing of hands and feet, fear of dying or going crazy) more intensely than their European counterparts (Friedman and Paradis 2002; Smith et al. 1999). In addition to somatic complaints, one phenomenon that has been found to be particularly salient in African Americans with PD is isolated sleep paralysis (Bell et al. 1986; Friedman and Paradis 1991, 2002; Friedman et al. 1994; Hinton et al. 2005; Otto et al. 2006; Paradis and Friedman 2005; Paradis et al. 1997). During isolated sleep paralysis, one may experience feelings of uncontrollable immobility, hallucinations, and feelings of impending danger, typically while in the process of waking or falling asleep. While the exact reason for these differences in rates of ISP remains relatively unknown, it has been hypothesized that they may be attributed to differences in levels of stress, sleep disruption, and hypertension.

Overall, the observed differential patterns of PD in African Americans may be associated with culturally specific variables and experiences. It has been hypothesized that the salience of somatic symptoms may be related to the fear of physical conditions (e.g., cardiovascular disease) and the underreporting of cognitive complaints resulting from stigma towards mental illness in African American communities (Hunter and Schmidt 2010 Carter et al. 1999; Friedman and Paradis 2002; Hunter and Schmidt 2010; Gordon and Teachman 2008; Johnson et al. 2007; Neal et al. 1994; Neal and Turner 1991). Additionally, exposure to chronic stress has been particularly implicated in PD with agoraphobia for African Americans (Barlow 1988; Carter and Barlow 1995). For example, the experience of chronic stress has been hypothesized to contribute to differences in symptom manifestation, including increased symptom severity, in African Americans. Therefore, the assessment of PD in this population must be sensitive to these cultural factors and the potential differences in symptom manifestation. In order to provide guidelines for culturally sensitive assessment of PD in African Americans, the "gold standard" assessment practices will be presented, as well as information regarding the utilization of current assessment tools with African Americans (e.g., psychometric properties) and cultural considerations in their use.

Assessment in Adult Populations

Interviewer Rating Scales

ADIS-IV Within the assessment of PD, the ADIS-IV measures the frequency, intensity, and duration of panic attacks, the antecedents to panic attacks, and avoidance behaviors associated with attacks. The ADIS-IV also assists the clinician to assess the history of the disorder, elucidate potential variations in manifestation patterns, and identify panic attacks that are un-cued, or unexpected (Brown et al. 1994). The PD section of the ADIS-IV has demonstrated sound psychometric properties. In a study examining the diagnostic reliability of the ADIS-IV in a sample of 363 individuals with anxiety disorders, interrater agreement of PD and PDA demonstrated a strong kappa ($\kappa = 0.79$; Brown et al. 2001). Additionally, the dimensional ratings regarding the frequency of panic attacks, fear of future attacks, agoraphobic avoidance, and clinical severity demonstrated good interrater reliability (alpha = 0.58–0.83). Although the ADIS-IV demonstrated sound psychometric properties in the aforementioned study, it is important to note that the study sample was predominately non-Hispanic White (e.g., 4 % African American). In a study that examined the efficacy of cognitive-behavioral treatment of PD and PDA in African Americans and European Americans, the ADIS-IV was used to determine a PD or PDA diagnosis (Friedman et al. 2006). It was found that African American participants were given a primary diagnosis PDA (rather than PD alone) significantly more than their European American counterparts. However, whether this finding reflected a true difference or a measurement error was not investigated. With regard to reliability, the ADIS-IV has been found to demonstrate excellent interrater reliability in exclusively African American samples (Chapman et al. under review). However, the reliability of the PD and PDA scales and validity of the ADIS-IV has yet to be examined in an exclusively African American sample, and possible differences in its utility with African Americans has not been explored.

Overall, the ADIS-IV remains a "gold standard" tool in the assessment of PD and PDA due to its empirically supported psychometric properties and its thorough assessment of the diagnostic criteria of the disorder. Despite a lack of study examining the reliability and validity of the ADIS-IV within ethnically diverse samples, the interview maintains important strengths in the assessment of PD and PDA in African Americans, such as providing a structural template for examining the full range of DSM-IV criteria as well as assessing specific elements of these disorders that are often underreported in African Americans (e.g., behavioral avoidance, panic-related cognitions).

Self-Report Measures

Panic Disorder Severity Scale (PDSS; Shear et al. 1997) The PDSS is a 7-item, clinician-rated measure utilized to assess multiple domains of panic and agoraphobia over the course of the past month. Specifically, the clinician indicates the frequency of panic attacks, the level of distress associated with panic attacks, fear of future attacks, avoidance of bodily sensations associated with panic, situational avoidance, and the impairment of panic on the client's social and occupational functioning. Ratings are measured on a 5-point Likert and the scales are average to create a composite score. The PDSS has consistently demonstrated sound reliability and internal consistency across samples (Shear et al. 1997, 2001b). The PDSS has been found to have convergent validity with clinical severity ratings on the ADIS-V, the corresponding items on the ADIS-V (e.g., fear of future attacks), and other self-report questionnaires of panic symptoms (e.g., Albany Panic and Phobia Questionnaires, Panic Disorder Self-Report). In a sample of 104 psychiatric outpatient participants, the PDSS was found to have an optimal cutoff score of 8 with a sensitivity of 83.3 % (Shear et al. 2001b).

Although the PDSS has demonstrated strong psychometric properties, the aforementioned reliability and validity studies did not include the ethnicities of their samples. As such, the psychometric properties and the normative data for this measure remain unknown in African American samples. In addition, due to a tendency to report somatic symptoms, African Americans may be rated inaccurately high on the scales that assess physical sensations and complaints. Moreover, African Americans may endorse more severe symptoms of PD, which may inflate the composite score and render an inaccurate representation of the severity of overall panic. Due to these potential confounds, it is suggested that the PDSS not be used in isolation in the assessment of the presence and severity of panic and agoraphobia symptoms, and that the aforementioned cutoff scores be referenced with caution.

Agoraphobic Cognitions Questionnaire (AgCQ; Chambless et al. 1984) The AgCQ is a 14-item self-report measure that was created to assess catastrophic cognitions related to the negative implications of anxiety, or the *fear of fear*. Clients indicate the frequency with which they experience the listed thought while anxious on a 5-point Likert scale (i.e., 1 = "thought never occurs"; 5 = "thought always occurs"). The questionnaire contains six items related to behavioral or social implications of anxiety and eight items pertaining to cognitions about physiological consequences. The AgCO has demonstrated a good internal consistency (Cronbach's alpha = 0.80) and a strong test-retest reliability (0.86). Additionally, the AgCO has been found to have convergent validity with other measures of agoraphobic cognitions and interoceptive cues, as well as discriminant validity in the delineation of agoraphobics and healthy controls (Chambless et al. 1984).

Overall, while evidence has supported the psychometric properties of the AgCO, the reliability and validity of this measure have not been examined in an exclusively African American sample. In addition, the extant literature on the AgCO does not include the ethnic demographics of their samples; therefore, it is unknown whether the current findings generalize to African American populations. Despite these limitations, one potential strength of the AgCO is its assessment of symptoms that may not be provided through verbal report by African American clients. Therefore, the AgCO may be a useful measure to gather information regarding cognitive symptoms of panic and agoraphobia, yet the yielded scores should be interpreted with caution.

Anxiety Sensitivity Index (ASI; Reiss et al. 1986) The ASI was developed to assess the construct of anxiety sensitivity, or the fear of the experiences related to anxiety (e.g., cognitive, physiological, evaluative). Notably, evidence supports that anxiety sensitivity predicts the development and maintenance of panic and related disorders (Maller and Reiss 1992; Schmidt et al. 1997; Taylor and Cox 1998). The ASI assesses one's level of anxiety sensitivity with 16 items that are rated on a 5-point Likert scale (i.e., 0 = "very little" to 4 = "very much"). Ratings reflect the degree to which one typically fears specific symptoms of anxiety. With regard to psychometrics, the ASI has demonstrated excellent internal consistency (alphas = 0.82–0.91) and good test-retest reliability ($r = 0.74$). For normative data, refer to Peterson and Reiss (1992). However, the aforementioned psychometric properties and norms do not include information on African American samples. Several studies have supported a three-factor structure of the ASI across diverse populations (e.g., outpatient, geriatric) including the factors of physical concerns, mental incapacitation concerns, and social concerns (Mohlman and Zinbarg 2000; Zinbarg et al. 1999). However, Carter et al. (1999) found that this three-factor structure did not demonstrate good fit with the data, and instead a four-factor model resulted in the best fit for a sample of 221 African American college students. This racially specific factor model included mental incapacitation, fears of unsteadiness, cardiovascular fears, and fears of losing emotional control, but aside from the mental incapacitation factor, this model did not demonstrate strong convergent validity with other anxiety measures. Furthermore, Arnau, Broman-Fulks, Green, and Berman (2009) examined the factor structures of the ASI and a revised version of the ASI (ASI-R; Taylor and Cox 1998) in a sample of African American and European American college students, and found that a separate four-factor structure (i.e., fear of respiratory symptoms, fear of publicly observable anxiety reactions, fear of cardiovascular symptoms, and fear of cognitive dyscontrol) delineated by Taylor and Cox (1998) provided the best fit across the two measures. Comparative analyses indicated that this four-factor model provided the best fit for the data within the African American sample, with the exception of one item that loaded on multiple factors (see Arnau et al. 2009). Further, an investigation of the reliability and validity of the ASI and ASI-R yielded strong internal consistencies. However, the ASI-R (Taylor and Cox 1998) demonstrated greater predictive ability as evidenced by greater correlations with both self-reported anxiety and specific diagnoses, including PD. Although the ASI-R was purported to be the optimal measure of anxiety sensitivity, the reliability and validity analyses did not include analyses specific to African Americans. As such, the psychometric findings may not generalize to exclusively African American samples.

Brief Panic Disorder Screen (BPDS; Apfeldorf et al. 1994) The BPDS is derived from the Anxiety Sensitivity Questionnaire, and includes four items that are purported to predict the presence of or vulnerability to PD (for a review of the psychometric properties, see Apfeldorf et al. 1994). In a study aimed at the examination of the comparative reliability and validity of the BPDS in African Americans ($n=79$) and European Americans ($n=218$), Johnson et al. (2007) found significant differences in the psychometric properties between the two populations. Specifically, the BPDS demonstrated lower internal consistency in the African American sample compared to the European American sample. The internal consistency was significantly lower for African Americans in clinical (i.e., PD diagnosis) and nonclinical subsets of the sample. In addition to decreased reliability, Johnson and colleagues found that the BPDS was significantly less accurate in predicting a PD diagnosis in African Americans as compared to European Americans. Overall, these findings suggest that the BPDS may not be appropriate in the assessment and determination of PD and PDA in African Americans. Moreover, these findings paired with the discrepant factor structure of the ASI suggest that the construct of anxiety sensitivity may differ in African American populations. As such, clinicians should be mindful of using current conceptualizations and measures of anxiety sensitivity to predict panic and related outcomes in African Americans. For further review of widely used measures of interoceptive fears related to PD that have not been evaluated in African Americans, see Chambless et al. (1984), Clark et al. (1997), and Austin et al. (2006).

Albany Panic & Phobia Questionnaire (APPQ; Rapee et al. 1995) The APPQ is a 27-item self-report measure that is used to assess aspects of agoraphobia and panic, including interoceptive fears and fears associated with agoraphobic and social phobic situations. Participants indicate the degree of fear they would expect to experience if they were confronted by the event in the following week on an 8-point Likert-type scale (0 = "no fear" to 8 = "extreme fear"). The APPQ yields three subscales, including Agoraphobia, Social Phobia, and Interoceptive Fears Scales. Research has supported the reliability and validity of the APPQ, with the measure demonstrating strong internal consistency, test-retest reliability, and construct validity (see Rapee et al. 1995). In addition, Chapman et al. (under review) found that the APPQ demonstrated high internal consistency in an exclusively African American adult sample (total $\alpha=0.86$). In a psychometric reevaluation of the APPQ, Brown, White, and Barlow (2005) confirmed the three-factor structure (i.e., agoraphobia, social phobia, and interoceptive) with all but three items loading on the indicated factor. Further, the factor scales demonstrated good internal reliability (alphas = 0.85–0.89), convergent validity, and divergent validity. Although the initial validation and reevaluation of the APPQ yield strong psychometric properties, the measure has yet to be examined in an African American sample within the context of PD. The sample in Brown and colleague's study included a small proportion of African Americans (i.e., 3.5 %), but the analyses employed did not investigate possible differences in the factor structure of the APPQ within this population. In addition, the clinical utility of the APPQ in the prediction of PD is warranted in African American samples. Overall, the APPQ appears to be an internally reliable measure in the assessment of panic and phobia in African Americans. While the scales of the APPQ may accurately reflect constructs related to panic and agoraphobia, optimal cut scores in the prediction of PD in African Americans have yet to be determined. Therefore, the interpretation of scale scores should be conducted in conjunction with other measures that assess the full range of diagnostic criteria of panic disorder with or without agoraphobia.

Agoraphobia

In addition to the items and subscales that assess the presence of agoraphobia and related constructs (e.g., agoraphobic cognitions) on the PDSS (Shear et al. 1997), AgCQ (Chambless et al. 1984), and APPQ (Rapee et al. 1995), several commonly used measures in the assessment of situational avoidance include the Mobility Inventory for Agoraphobia (Chambless et al. 1985) and The Fear Questionnaire (Marks and Matthews 1979). Currently, the aforementioned measures have yet to be examined in African American samples, and will therefore not be discussed. For a review of the Mobility Inventory for Agoraphobia and the Fear Questionnaire, please see Chambless et al. (1985) and Marks and Matthews (1979). When using these measures and assessing for the presence of agoraphobia in African Americans, one should interpret the validity of scale scores with caution. Additionally, self-report measures should not be used in isolation as indicators of agoraphobia, and are best interpreted in conjunction with assessment measures and tools that assess the full range of diagnostic criteria of agoraphobia. Particular attention should also be given to ensure that agoraphobia assessment batteries include measures of symptoms that are often underreported in African American samples, such as cognitive symptoms. In terms of assessment of children, several assessment tools that measure constructs related to PD in children have been examined in African American children, including the Screen for Child Anxiety Related Emotional Disorders-Panic (SCARED-P) and Child Anxiety Sensitivity Index (CASI).

Screen for Child Anxiety Related Emotional Disorders-Panic Subscale (SCARED; Birmaher et al. 1999) The SCARED is a 41-item self-report measure used to assess various anxiety disorders in children, including panic, GAD, separation anxiety, social phobia, and school phobia. The panic subscale consists of 13 items rated on 3-point Likert scale that indicate the extent to which a child experiences symptoms of PD (e.g., "When I am frightened, it is hard to breathe.") Original psychometric analyses revealed that the SCARED demonstrates good internal consistency (coefficients = 0.74–0.89), test-retest reliability (0.70–0.90), and validity in the discrimination of clinical and nonclinical youth (Birmaher et al. 1997). Two extant studies specifically examined the reliability of the SCARED in samples of African American youth, and obtained good internal consistencies of the panic subscale that were comparable to those of the original psychometric evaluation (0.74–0.88; Ginsburg et al. 2004; Gonzalez et al. 2012). In addition, in a sample of 64 African American youth between the ages of 14 and 19, the panic subscale demonstrated a 6-month test-retest reliability of 0.40.

Gonzalez et al. (2012) evaluated the measurement equivalence of the parent and child versions of the SCARED in a comparative sample of outpatient African American and non-Hispanic White youth and their parents. Descriptive fit indices indicated that the original five-factor model of the SCARED-C demonstrated adequate fit within the African American sample. In terms of individual factor loadings, it was found that one item (i.e., "When my child gets frightened, he/she feels like passing out") loaded more strongly on the Panic/Somatic scale for African Americans based on both parent and child reports. In addition to the factor structure of the SCARED, the study examined the concurrent criterion validity of the parent-version of the SCARED. Results indicated that the total cutoff score of 25 was a significant predictor of the presence of an anxiety disorder, with a sensitivity of 60 % and specificity of 88 % in the African American sample. However, the results failed to looked at the predictive validity of the individual subscales to their respective diagnoses. Therefore, the SCARED-P was only supported as a screener of an anxiety disorder in African American youth.

Overall, burgeoning research suggests that the SCARED (both parent and child versions) is an internally consistent, reliable measure of anxiety in African American youth. Specifically, it appears as though the panic subscale of the measure demonstrates reliability in samples of African American children. However, the validity of the panic subscale, including its ability to predict clinical levels of panic, needs to be examined in African American youth. Therefore, the SCARED is recommended as an appropriate screener for the presence of an anxiety disorder and panic symptoms in African American children.

Child Anxiety Sensitivity Index (CASI; Silverman et al. 1991) The CASI is an 18-item self-report measure that is a modified version of the Anxiety Sensitivity Index for children that is utilized to assess the extent to which one fears the symptoms associated with anxiety. The items are rated on a 3-point Likert scale (0 = "a lot" to 3 = "a lot") and represent the level of fear towards a specific symptom. The individual responses are summed to obtain a total score of anxiety sensitivity. The psychometric properties of the CASI, including internal consistency, test-retest reliability, and construct validity, have been supported in both community and clinical samples (Chorpita et al. 1998; Rabian et al. 1999; Silverman et al. 1991; Weems et al. 1998).

Recent work has also examined the psychometric properties of the CASI in exclusively African American samples. Specifically, Lambert et al. (2004) examined the reliability, validity, and factor structure of the CASI in a sample of 144 elementary-aged African American children. The results yielded comparable internal consistencies to the original psychometric evaluation (Lambert et al. 2004); however, African American children demonstrated high mean levels of anxiety sensitivity than means previously identified in non-Hispanic White youth (for means, refer to Lambert et al. 2004). Further, exploratory and confirmatory factor analyses resulted in a two-factor structure (i.e., physical concerns and mental incapacitation) that was discrepent from Previous factor structures found in non-Hispanic White samples (Silverman et al. 1999). Utilizing the two-factor subscales and total scores, the CASI demonstrated convergent validity with measures of general, physiological, and social anxiety. Moreover, divergent validity was established with scores of perfectionism, anxious coping, and depression. The validity of the CASI as it relates to PD in African American youth has also received initial support, as Ginsburg and Drake (2002) found significant correlations between the CASI and Panic Attack Questionnaire (PAQ; Norton et al. 1986).

Overall, there is initial support for the use of the CASI in the assessment of anxiety sensitivity in African American youth. In addition, burgeoning research has provided preliminary evidence for the relationship of anxiety sensitivity to PD in this population. While the CASI appears to have adequate psychometric properties in African American samples, the measure may not fully capture the nature of this construct within African American youth.

Specific Phobia

The presence of a specific phobia is characterized by an excessive fear of an explicit situation or object that is persistent in nature. Further, this fear results in the avoidance of the specified stimuli or significant distress when the phobic target is encountered (American Psychiatric Association 2000). Exposure to the phobic object results in an immediate anxiety response, which may include the experience of a panic attack. Current categorizations of this disorder delineate phobias into five subtypes, including animal, natural environment, blood-injection-injury, situational, and other (e.g., fears not captured by the aforementioned categories) fears.

The existing research suggests that African American adults experience comparable to increased rates of specific phobia compared to non-Hispanic Whites (Breslau et al. 2006; Brown et al. 1990; Last and Perrin 1993; Neal and Turner 1991; Warheit et al. 1975). Specifically, extant studies examining the presence of specific phobia in African Americans adults have found endorsement rates up to three times greater than European Americans, even when demographic variables (e.g., SES, education, geographic location) are controlled (Last and Perrin 1993; Neal and Turner 1991). Increased prevalence rates of specific phobia have also been demonstrated in African American children. For example, in a study of 162 African American and non-Hispanic White children with anxiety disorders, African American children demonstrated higher incidence rates (i.e., 26.7 %) and lifetime prevalence rates (i.e., 50 %) than their non-Hispanic White counterparts. One cultural factor that may contribute to these observed differences is the exposure to high stress environments. Specifically, chronic stress may increase the experience of acute and chronic fearful states, which has been theorized to account for the elevated rates of specific phobia in this population (Carter et al. 1996; Neal and Turner 1991). In addition to variation in observed prevalence rates of specific phobia, the existing literature suggests that African Americans demonstrate differences in the content of fears. Studies examining fear content in African American adults and children have highlighted the tendency for this population to endorse more "reality-based" fears than their non-Hispanic White counterparts. For example, African American adults have been found to report more animal fears (Chapman et al. 2008, 2011, 2012; Nalven 1970) and fears associated with natural or environmental conditions (e.g., water; Chapman et al. 2008). Research suggests that African American children also endorse more specific animal fears and an overall greater number of fears than non-Hispanic White children (LaPouse and Monk 1959; Nalven 1970; Neal et al. 1993). While these differences in fear content have been observed, the studies examining these constructs have often investigated fear in a nonpathological context. Therefore, future research is needed to understand whether these differential fear patterns pervade clinical diagnoses, as well as the cultural factors that may influence their etiology and expression.

Interviewer Rating Scales

ADIS-IV Within the assessment of specific phobias, the ADIS-IV (Brown et al. 1994) obtains the severity of fear and avoidance of specific stimuli, as well as the symptomatic behaviors associated with exposure to the phobic stimuli (e.g., panic attacks, automatic anxiety response). Further, in order to determine a clinical diagnosis, the ADIS-IV assesses the functional impairment resulting from the specific phobia.

The specific phobia section of the ADIS-IV has consistently demonstrated strong psychometric properties, including good reliability. Chapman et al. (under review) found that the ADIS-IV demonstrated excellent inter-rater reliability in a sample of 65 African American adults; however, analyses did not examine reliability exclusively in the context of specific phobia. In a study examining interrater reliability of ADIS-IV diagnoses in a sample of 362 adults, the kappa for the specific phobia section was 0.86 (Brown et al. 2001). Although the ADIS-IV has been deemed a psychometrically sound measure for the assessment of specific phobia in predominately non-Hispanic White samples, its validity has not been examined in African American samples. For example, only 4 % of the sample in the Brown et al. (2001) study self-identified as African American. As such, empirical investigation is warranted to explore whether the psychometric properties of the specific phobia scale are maintained in African American populations. Overall, the ADIS-IV demonstrates strength in its assessment of a full range of symptoms and behaviors associated with specific phobia. Moreover, its semistructured nature allows for some clinician flexibility to prompt potential cultural factors that impact a client's specific phobia development, manifestation, or maintenance.

Self-Report Scales

Fear Survey Schedule-Second Edition (FSS-II; Geer 1965) The FSS-II is one of the most commonly used self-report screening measures used in the assessment of specific phobia. The measure is comprised of 51 items that assess the fear associated with common specific phobia situations, social phobia situations, phobic objects, and agoraphobia. Factor analytic studies have indicated that that FSS-II includes the factors of water, death, illness and injury, concrete objects, live organisms, violence, social interaction, and negative social evaluation (Bernstein and Allen 1969; Rubin et al. 1968). The FSS-II has demonstrated strong internal reliability ($r = 0.94$; Geer 1965) and validity; however, these analyses were conducted on predominately European American samples and may not generalize to African Americans.

Recent research has examined the factor structure of the FSS-II in both college and community samples of African Americans. Specifically, Chapman et al. (2008) utilized confirmatory factor analysis to evaluated the factor structure of the FSS-II in both African American and European American college students. Results indicated differential factor structures and patterns of endorsed fears between the two samples. A three-factor model comprised of natural environment, animal, and social anxiety factors emerged within the African American sample. In addition, African Americans endorsed more overall fears and fears related to animals and the natural environment than their European American counterparts. Cross-validation of the FSS-II in a community sample of African Americans corroborated these differential patterns of fear endorsement (Chapman et al. 2011). In particular, the factors of animal and social fears demonstrated consistent endorsement in African American adults; however, the community sample did not endorse the same extent of natural environment fears as the college sample.

Overall, the extant research on the FSS-II suggests that African Americans may endorse different patterns of fears than European Americans, resulting in a differential factor structure of the measure. While the factor structure of this measure has been examined in both college and community samples, its psychometric properties have yet to be determined within African Americans. In addition, the FSS-II has traditionally been utilized as a measure of fear for research purposes, and its clinical utility is unknown. A subsequent version of the measure, the Fear Survey Schedule-Third Edition (FSS-III; Wolpe and Lange 1964), has been utilized as a clinical screen for phobic symptoms. However, its discriminative validity has not been supported (Klieger and Franklin 1993) and it has not been examined in African American samples.

Revised Fear Survey Schedule for Children (FSSC-R; Ollendick 1983) The FSSC-R is an 80-item self-report measure that assesses common fears in children. The items are measured on a three-point Likert scale (i.e., “none” to “a lot”) to indicate the level of fearfulness of the situation or object. The FSSC-R has demonstrated strong reliability and validity (Ollendick 1983), and has been widely used in both research and clinical populations (King et al. 1989, 1990; Ollendick et al. 1989, 1991). Factor analyses of the FSSC-R have resulted in a five-factor model, including fear of failure and criticism, fear of the unknown, fear of injury and small animals, fear of danger and death, and medical fears (Ollendick 1983; Ollendick et al. 1985). With regard to clinical utility, the FSSC-R has been found to differentiate nonclinical and school-phobic children, and has been suggested for use as a normative screen for fear sensitivities in children (Ollendick 1983).

Although the FSSC-R has demonstrated consistent reliability across studies and initial support for its clinical utility, a paucity of studies have evaluated this measure within African American populations. One extant study examined the factor structure of the FSSC-R and stability of child fear in a comparative sample of African American and non-Hispanic White children (Neal et al. 1993). Within this study, the FSSC-R demonstrated excellent internal consistency within the African American sample (cronbach’s alpha = 0.96). Analysis of the stability of children’s reported fears over a two-week period yielded less stability in African American children than their European American counterparts (stability coefficients = 0.71, 0.83). In addition, the FSSC-R was found to demonstrate a differential factor structure between the two samples. Specifically, a three-factor model including fear of death, danger, and small animals, fear of the unknown and things that crawl, and medical fears. Although the FSSC-R yielded discordant factor structures for African American children and European American children, there was significant overlap in the most frequently endorsed fears between these two populations.

Overall, the FSSC-R has received initial support for its reliability within African American populations. However, additional studies are needed to confirm the factor structure yielded in Neal et al. (1993) and if the FSSC-R needs to be modified to more accurately capture the content of African American children’s fear. In addition, research evaluating the clinical utility of the FSSC-R in the assessment of specific phobias in African American youth is warranted. As such, the FSSC-R may be useful in the assessment of the content of a child's fear, but not in the determination of whether this fear is pathological in nature within African American populations.

Summary and General Recommendations

The main objective of this chapter was to provide a framework for the assessment of generalized anxiety disorder, social phobia, panic disorder, and specific phobias in African Americans. This review of the extant literature highlighted the burgeoning nature of our understanding of the aforementioned disorders and their assessment. Specifically, several individual "gold standard" assessment tools reviewed have been explored in exclusively African American samples. Despite these significant strides, continued research is warranted to ensure that individual assessment tools and comprehensive evaluations fully capture the diagnostic presentation of these anxiety disorders in African Americans while being sensitive to the variation within this population. In particular, several assessment tools (e.g., ADIS-IV, SPAI, HAS, PDSS, AcCO) have yet to be validated for use in African American samples. Additionally, the replication of the psychometric properties and exploration of the clinical utility of other gold standard measures (e.g., APPQ, PSWQ, WDQ, FSSC) is warranted. Future empirical investigation is also needed to account for the heterogeneity in African Americans when assessing these disorders in this population. For example, research should examine the impact of various sociocultural variables (e.g., ethnic identity, acculturation) on the psychometric and clinical utility of "gold standard" measurement tools.

While continued empirical investigation is warranted, practitioners and researchers can take steps to ensure that they are employing culturally-sensitive assessment practices until a more thorough research base has been established. In particular, when assessing GAD, PD, and the phobias in African American samples, it is imperative for one to understand the individual's context and the sociocultural variables (e.g., ethnic identity, acculturation, socioeconomic status, education background) that may impact the manifestation and explanatory models of their anxiety (e.g., "meet the client where they're at"). In addition, the following recommendations are made to ensure "gold standard" practices that fully capture the symptom presentation of GAD, PD, and the phobias in African American individuals: (1) Avoid using any assessment tool in isolation that has yet to be validated in African American samples; (2) Examine individual item responses when scale norms have not been established or fully replicated; (3) Ensure that assessment practices are examining the full-range of diagnostic criteria, with particular care to assess for symptoms that may be traditionally underreported in African American samples (e.g., cognitive symptoms of anxiety).

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders—text revision* (4th ed.). Washington, DC: American Psychiatric Association.

Apfeldorf, W., Shear, M. K., Leon, A. C., & Portera, L. (1994). A brief screen for panic disorder. *Journal of Anxiety Disorders, 8,* 71–78.

Arnau, R. C., Broman-Fulks, J., Green, B., & Berman, M. (2009). The anxiety sensitivity index-revised: Confirmatory factor analyses, structural invariance in Caucasian and African American samples, and score reliability and validity. *Assessment, 16*(2), 165–180.

Austin, D. W., Richards, J. C., & Klein, B. (2006). Modification of the body sensations interpretation questionnaire (BSIQ-M): Validity and reliability. *Journal of Anxiety Disorders, 20*(2), 237–251.

Baker, S. L., Hendrichs, N., Kim, H. J., & Hofmann, S. G. (2002). The Liebowitz social anxiety scale as a self-report instrument: a preliminary psychometric analysis. *Behaviour Research and Therapy, 40,* 701–715.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford.

Barlow, D. H. (2004). Anxiety and its disorders: *The nature and treatment of anxiety and panic*. Guilford press.

Beard, C., Rodriguez, B. F., Moitra, E., Sibrava, N. J., Bjornsson, A., Weisberg, R. B., & Keller, M. B. (2011). Psychometric properties of the Liebowitz social anxiety scale (LSAS) in a longitudinal study of African Americans with anxiety disorders. *Journal of Anxiety Disorders, 25,* 722–726.

Beck, A. T., & Steer, R. A. (1987). *Beck depression inventory manual*. San Antonio: The Psychological Corporation.

Beidel, D. C., Turner, S. M., & Morris, T. L. (1995). A new inventory to assess childhood social anxiety and phobia: The social and anxiety inventory for children. *Psychological Assessment, 7,* 73–79.

Beidel, D., Turner, S. M., Hamlin, K., & Morris, T. L. (2000). The social phobia and anxiety inventory for children (SPAI-C): External and discriminant validity. *Behavior Therapy, 31,* 75–87.

Bell, C. C., Dixie-Bell, D. D., & Thompson, B. (1986). Further studies on the prevalence of isolated sleep paralysis in Black subjects. *Journal of the National Medical Association, 78,* 649–659.

Bernstein, D. A., & Allen, G. J. (1969). Case histories and shorter communications. *Behavior Research and Therapy, 7,* 403–407.

Birmaher, B., Khetarpal, S., Brent, D., & Cully, M. (1997). The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry, 36,* 545–553.

Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the screen for child anxiety related emotional disorders (SCARED): A replication study. *Journal of the American Academy of Child & Adolescent Psychiatry, 38,* 1230–1236.

Breslau, J., Aguilar-Gaxiola, S., Kendler, K. S., Su, M., Williams, D., & Kessler, R. C. (2006). Specifying raceethnic differences in risk for psychiatric disorder in a USA national sample. *Psychological Medicine, 36,* 57–68.

Brown, D. R., & Eaton, W. W. (1986). *Racial differences in risk factors for phobic disorders*. Paper presented at the 114th meeting of the American Public Health Association.

Brown, D. R., Eaton, W., & Sussman, L. (1990). Racial differences in prevalence of phobic disorders. *Journal of Nervous & Mental Disease, 178*(7), 434–441.

Brown, T. A., Di Nardo, P. A., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV*. New York: Oxford University Press.

Brown, T. A., Di Nardo, P. A., Lehman, C. L., & Campbell, L. A. (2001). Reliability of *DSM-IV* anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology, 110*(1), 49–58.

Brown, T. A., White, K. S., & Barlow, D. H. (2005). A psychometric reanalysis of the Albany Panic and Phobia Questionnaire. *Behaviour Research and Therapy, 43,* 337–355.

Bruss, G. S., Gruenberg, A. M., Goldstein, R. D., & Barber, J. P. (1994). Hamilton anxiety rating scale interview guide: Joint interview and test-retest methods for interrater reliability. *Psychiatry Research, 53*(2), 191–202.

Carter, M. M., & Barlow, D. H. (1995). Learned alarms: The origins of panic. In W. O'Donohue & L. Krasner (Eds.), *Theories of behavior therapy: Exploring behavior change*. Washington, DC: APA Books.

Carter, M. M., Sbrocco, T., & Carter, C. (1996). African Americans and anxiety disorders research: Development of a testable theoretical framework. *Psychotherapy, 33*(3), 449–463.

Carter, M. M., Miller, O., Sbrocco, T., Suchday, S., & Lewis, E. L. (1999). Factor structure of the anxiety sensitivity index among African American students. *Psychological Assessment, 11,* 525–533.

Carter, M. M., Sbracco, T., Miller, O., Suchday, S., Lewis, E. L., & Freedman, R. E. K. (2005). Factor structure, reliability, and validity of the Penn State worry questionnaire: Differences between African-American and White-American college student. *Journal of Anxiety Disorders, 19,* 827–843.

Chambless, D. L., Caputo, C., Bright, P., & Gallagher, R. (1984). Assessment of fear of fear in agoraphobics: The body sensations questionnaire and the agoraphobic cognitions questionnaire. *Journal of Consulting and Clinical Psychology, 52*(6), 1090–1097.

Chambless, D. L., Caputo, G. C., Jasin, S., Gracely, E., & Williams, C. (1985). The mobility inventory for agoraphobia. *Behaviour Research and Therapy, 23*(1), 35–44.

Chapman, L. K., Petrie, J. M., & Richards, A. E. (under review). The efficacy of self-report measures in predicting social phobia in African American adults.

Chapman, L. K., Kertz, S. J., & Zurlage, M., & Woodruff-Borden, J. (2008). A confirmatory factor analysis of specific phobia domains in African American and Caucasian American young adults. *Journal of Anxiety Disorders, 22*(5), 763–771.

Chapman, L. K., Kertz, S. J., & Woodruff-Borden, J. (2009). A structural equation model analysis of perceived control and psychological distress on worry among African American and European American young adults. *Journal of Anxiety Disorders, 23,* 69–76.

Chapman, L. K., Vines, L., & Petrie, J. (2011). Fear factors: Cross validation of specific phobia domains in a community-based sample of African American adults. *Journal of Anxiety Disorders, 25*(4), 539–544.

Chapman, L. K., Petrie, J., Vines, L., & Durrett, E. (2012). The co-occurrence of anxiety disorders in African American parents and their children. *Journal of Anxiety Disorders, 26*(1), 65–70.

Chorpita, B. F., Tracey, S. A., Brown, T. A., Collica, T. J., & Barlow, D. H. (1997). Assessment of worry in children and adolescents: An adaptation of the penn state worry questionnaire. *Behavior Research and Therapy, 35,* 569–581.

Chorpita, B. F., Brown, T. A., & Barlow, D. H. (1998). Perceived control as a mediator of family environment in etiological models of childhood anxiety. *Behavior Therapy, 29,* 457–476.

Clark, D. M., Salkovskis, P. M., Ost, L. G., Breitholtz, E., Koehler, K., Westling, B. E., Jeavons A., & Gelder M. (1997). Misinterpretation of body sensations in panic disorder. *Journal of Consulting and Clinical Psychology, 65*(2), 203–213.

Ferrell, C. B., Beidel, D. C., & Turner, S. M. (2004). Assessment and treatment of socially phobic children: A cross cultural comparison. *Journal of Clinical and Adolescent Psychology, 33*(2), 260–268.

Friedman, S., & Paradis, C. (1991). African-American patients with panic disorders and agoraphobia. *Journal of Anxiety Disorders, 5,* 35–41.

Friedman, S., & Paradis, C. (2002). Panic disorder in African-Americans: Symptomatology and isolated sleep paralysis. *Culture, Medicine, and Psychiatry, 26,* 179–198.

Friedman, S., Paradis, C., & Hatch, M. (1994). Characteristics of African American and White patients with panic disorder and agoraphobia. *Hospital and Community Psychiatry, 45,* 798–803.

Friedman, S., Braunstein, J. W., & Halpern, B. (2006). Cognitive behavioral treatment of panic disorder and agoraphobia in a multiethnic urban outpatient clinic: Initial presentation and treatment outcome. *Cognitive and Behavioral Practice, 13*(4), 282–292.

Geer, J. H. (1965). The development of a scale to measure specific phobia. *Behavior Research and Therapy, 3,* 45–53.

Ginsburg, G. S., & Drake, K. L. (2002). Anxiety sensitivity and panic attack symptomatology among low-income African-American adolescents. *Journal of Anxiety Disorders, 16*(1), 83–96.

Ginsburg, G. S., Lambert, S. F., & Drake, K. L. (2004). Attributions of control, anxiety sensitivity, and panic symptoms among adolescents. *Cognitive Therapy & Research, 28*(6), 745–763.

Gonzalez, A., Weersing, V., Warnick, E., Scahill, L., & Woolston, J. (2012). Cross-ethnic measurement equivalence of the SCARED in an outpatient sample of African American and Non-Hispanic White youths and parents. *Journal of Clinical Child and Adolescent Psychology, 41*(3), 361–369.

Gordon, T. L., & Teachman, B. A. (2008). Ethnic group differences incaffective, behavioral, and cognitive markers of anxiety. *Journal of Cross-Cultural Psychology, 39,* 424–446.

Hambrick, J. P., Rodebaugh, T. L., Balsis, S., Woods, C. M., Mendez, J. L., & Heimberg, R. G. (2010). Cross-ethnic measurement equivalence of measures of depression, social anxiety, and worry. *Assessment, 17*(2), 155–171.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32,* 50–55.

Heimberg, R. G., Horner, K. J., Juster, H. R., Safren, S. A., Brown, E. J., Schneier, F. R., et al. (1999). Psychometrics properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine, 29,* 199–212.

Himle, J. A., Baser, R. E., Taylor, R. J., Campbell, R. D., & Jackson, J. S. (2009). Anxiety disorders among African Americans, blacks of Caribbean descent, and non-Hispanic whites in the United States. *Journal of Anxiety Disorders, 23,* 578–590.

Hinton, D. E., Hufford, D. J., & Kirmayer, L. J. (2005). Culture and sleep paralysis. *Transcultural Psychiatry, 42,* 5–10.

Horwath, E., Johnson, J., & Hornig, C. D. (1994). Epidemiology of panic disorder. In S. Friedman (Ed.), *Anxiety disorders in African Americans* (pp. 53–64). New York: Springer.

Hughes, A. A., & Kendall, P. C. (2009). Psychometrics properties of the positive and negative affect scale for children (PANAS-C) in children with anxiety disorders. *Child Psychiatry and Human Development, 40*(3), 343–352.

Hunter, L. R., & Schmidt, N. B. (2010). Anxiety psychopathology in African American adults: Literature review and development of an empirically informed sociocultural model. *Psychological Bulletin, 136*(2), 211–235.

Johnson, M. R., Hartzema, A. G., Mills, T. L., De Leon, J. M., Yang, M., Frueh, C., & Santos, A. (2007). Ethnic differences in the reliability and validity of a panic disorder screen. *Ethnicity and Health, 12,* 283–296.

King, N. J., Ollier, K., Iacuone, R., Schuster, S., Bays, K., Gullone, E., & Ollendick, T. H. (1989). Fears of children and adolescents: A cross-sectional Australian study using the Revised-Fear Survey Schedule for Children. *Child Psychology 7 Psychiatry & Allied Disciplines, 30*(5), 775–784.

King, N. J., Gullone, E., & Ollendick, T. H. (1990). Fears in children and adolescents with chronic medical conditions. *Journal of Clinical Child Psychology, 19*(20), 173–177.

Klieger, D. M., & Franklin, M. E. (1993). Validity of the Fear Survey Schedule in phobia research: A laboratory test. *Journal of Psychopathology and Behavioral Assessment, 15*(3), 207–217.

Lambert, S. F., Cooley, M. R., Campbell, K. M., Benoit, M. Z., & Standbury, R. (2004). Assessing anxiety sensitivity in inner-city African American children: Psychometric properties of the childhood anxiety sensitivity index. *Journal of Clinical Child and Adolescent Psychology, 33*(2), 248–259.

Lapouse, R., & Monk, M. A. (1959). Fears and worries in a representative sample of children. *American Journal of Orthopsychiatry, 29,* 803–818.

Last, C. G., & Perrin, S. (1993). Anxiety disorders in African-American and white children. *Journal of Abnormal Child Psychology, 21*(2), 153–164.

Laurent, J., Catanzaro, S., Joiner, T., Rudolf, K., Potter, K., & Lambert, S. (1999) A measure of positive and negative affect for children: scale development and preliminary validation. *Psychological Assessment, 11*(3), 326–338.

Liebowitz, M. R. (1987). Social phobia. *Modern Problems in Pharmacopsychiatry, 22,* 141–173.

Luterek, J. A., Turk, C. L., Heimberg, R. G., Frescro, D. M., & Mennin, D. S. (2002). *Psychometric properties of the GAD-Q-IV among individuals with clinician-*

assessed generalized anxiety disorder: An update. Paper presented at the Association for Advancement of Behavior Therapy (Reno, NV) November 2002.

Maller, R. G., & Reiss, S. (1992). Anxiety sensitivity in 1984 and panic attacks in 1987. *Journal of Anxiety Disorders, 6*(3), 241–247.

Marks, I. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy, 17*(3), 263–267.

Mattick, R. P., & Clark, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy, 36,* 455–470.

Melka, S. E., Lancaster, S. L., Adams, L. J., Howarth, E. A., & Rodriguez, B. F. (2010). Social anxiety across ethnicity: A confirmatory factor analysis of the FNE and SAD. *Journal of Anxiety Disorders, 24,* 680–685.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State worry questionnaire. *Behaviour Research and Therapy, 28,* 487–495.

McLaughlin, K. A., Hilt, L. M., & Nolen-Hoekema, S. (2007). Racial/ethnic differences in internalizing and externalizing symptoms in adolescents. *Journal of Abnormal Child Psychology, 35,* 801–816.

Mohlman, J., & Zinbarg, R. E. (2000). The structure and correlates of anxiety sensitivity in older adults. *Psychological Assessment, 12*(4), 440–446.

Nalven, F. B. (1970). Manifest fears and worries of ghetto vs middle-class suburban children. *Psychological Reports, 27*(1), 285–290.

Neal, A. M., & Turner, S. M. (1991). Anxiety disorders research with African Americans: Current status. *Psychological Bulletin, 109,* 400–410.

Neal, A. M., Lilly, R. S., & Zakis, S. (1993). What are African American children afraid of. *Journal of Anxiety Disorders, 7,* 129–139.

Neal, A. M., Rich, L. N., & Smucker, W. D. (1994). The presence of panic disorder among African American hypertensives: A pilot study. *Journal of Black Psychology, 20,* 29–35.

Newman, M. G., Zuellig, A. R., Kachin, K. E., Constantino, M. J., & Cashman, L. (2002). The reliability and validity of the GAD-Q-IV: A revised self-report diagnostic measure of generalized anxiety disorder. *Behavior Therapy, 33,* 215–233.

Norton, G. R., Dorward, J., & Cox, B. J. (1986). Factors associated with panic attacks in non-clinical subjects. *Behavior Therapy, 17,* 239–252.

Ollendick, T. H. (1983). Reliability and validity of the revised fear survey schedule for children (FSSC-R). *Behaviour Research and Therapy, 21*(6), 685–692.

Ollendick, T. H., Matson, J. L., Helsel, W. J. (1985). Fears in children and adolescents: Normative data. *Behaviour Research and Therapy, 23*(4), 465–467.

Ollendick, T. H., King, N. J., & Frary, R. B. (1989). Fears in children and adolescents: Reliability and generalizability across gender, age, and nationality. *Behaviour Research & Therapy, 27*(1), 19–26.

Ollendick, T. H., Yule, W., & Ollier, K. (1991). Fears in British children and their relationship to manifest anxiety and depression. *Journal of Child Psychology and Psychiatry, 32,* 321–331.

Otto, M. W., Simon, N. M., Power, M., Hinton, D., Zalta, A. K., & Pollack, M. H. (2006). Rates of isolated sleep paralysis in outpatients with anxiety disorders. *Anxiety Disorders, 20,* 687–693.

Paradis, C. M., & Friedman, S. (2005). Sleep paralysis in African Americans with panic disorder. *Transcultural Psychiatry, 42,* 123–134.

Paradis, C. M., Friedman, S., Lazar, R. M., Grubea, J., & Kesselman, M. (1992). Use of a structured interview to diagnose anxiety disorders in a minority population. *Hospital & Community Psychiatry, 43*(1), 61–64.

Paradis, C. M., Hatch, M., & Friedman, S. (1993). Anxiety disorders in African American: An update. *Journal of the National Medical Association, 86*(8), 609–612.

Paradis, C. M., Friedman, S., & Hatch, M. J. (1997). Isolated sleep paralysis in African Americans with panic disorder. *Journal of Cultural Diversity and Mental Health, 3,* 69–76.

Petrie, J. M., Chapman, L. K., & Vines, L. M. (2013). Utility of the PANAS-X in predicting social phobia in African American females. *Journal of Black Psychology*, *39*(2), 131–155.

Peterson, R. A., & Reiss, S. (1992). *Anxiety sensitivity index manual* (2nd ed.). Worthington: International Diagnostic Systems.

Rabian, B., Embry, L., & MacIntyre, D. (1999). Behavioral validation of the Childhood Anxiety Sensitivity Index in children. *Journal of Clinical Child Psychology, 28,* 105–112.

Rapee, R. M., Craske, M. G., & Barlow, D. H. (1995). Assessment instrument for panic disorder that includes fear of sensation-producing activities: The Albany Panic and Phobia Questionnaire. *Anxiety, 1,* 114–122.

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency, and the prediction of fearfulness. *Behaviour Research and Therapy, 24,* 1–8.

Roberson-Nay, R., Strong, D. R., Nay, W. T., Beidel, D. C., & Turner, S. M. (2007) Development of an abbreviated Social Phobia and Anxiety Inventory (SPAI) using item response theory: The SPAI-23. *Psychological Assessment, 19*(1), 133–145.

Robinson, C. M., Klenck, S. C., & Norton, P. J. (2010). Psychometric properties of the Generalized Anxiety Disorder Questionnaire for DSM-IV among four racial groups. *Cognitive Behaviour Therapy, 39*(4), 251–261.

Rubin, B. M., Katkin, E. S., & Weiss, B. W. (1968). Factor analysis of a Fear Survey Schedule. *Behavior Research and Therapy, 6*(1), 65–76.

Rucker, L. S., West, L. M., & Roemer, L. (2010). Relationships among perceived racial stress, intolerance of uncertainty, and worry in a Black sample. *Behavior Therapy, 41,* 245–253.

Ruscio, A. M., Brown, T. A., Chiu, W. T., et al. (2008). Social fears and social phobia in the United States:

Results from the National Comorbidity Survey Replication. *Psychological Medicine, 38*(1), 15–28.

Schry, A. R., Roberson-Nay, R., & White, S. W. (2012). Measuring social anxiety in college students: A comprehensive evaluation of the psychometric properties of the SPAI−23. *Psychological Assessment, 24*(4), 846–854.

Safren, S. A., Heimberg, R. G., Horner, K. J., Juster, H. R., Schneier, F. R., & Liebowitz, M. R. (1999). Factor structure of social fears: the Liebowitz Social Anxiety Scale. *Journal of Anxiety Disorders, 13,* 253–270.

Schmidt, N. B., Larew, D. R., & Jackson, R. J. (1997). The role of anxiety sensitivity in the pathogenesis of panic: Prospective evaluation of spontaneous panic attacks during acute stress. *Journal of Abnormal Psychology, 106,* 355–364.

Scott, E. L., Eng, W., & Heimberg, R. G. (2002). Ethnic differences in worry in a nonclinical population. *Depression and Anxiety, 15,* 79–82.

Shear, M. K., Brown, T. A., Barlow, D. H., Money, R., Sholomskas, D., Woods, S. W., Gorman, J., & Papp, L. (1997). Multicenter collaborative panic disorder severity scale. *American Journal of Psychiatry, 154,* 1571–1575.

Shear, M. K., Bilt, J. V., Rucci, P., Endicott, J., Lydiard, B., Otto, M. W., Pollack, M. H., et al. (2001a). Reliability and validity of a structured interview guide for the Hamilton Anxiety Rating Scale (SIGH-A). *Depression and Anxiety, 13,* 166–178.

Shear, M. K., Rucci, P., Williams, J., Frank, E., Grochocinski, V., Vander Bilt, J., Houck, P., & Wang, T. (2001b). Reliability and validity of the Panic Disorder Severity Scale: Replication and extension. *Journal of Psychiatric Research, 35,* 293–296.

Silverman, W. K., & Albano, A. M. (1996). *The anxiety disorders interview schedule for children for DSM-IV: (Child and parent versions)*. San Antonio: Psychological Corporation.

Silverman, W. K., Fleisig, W., Rabian, B., & Peterson, R. A. (1991). Childhood anxiety sensitivity index. *Journal of Clinical Child Psychology, 20,* 162–168.

Silverman, K. W., Ginsburg, G. S., & Goedhart, A. W. (1999). Factor structure of the childhood anxiety sensitivity index. *Behaviour Research and Therapy, 37,* 903–917.

Smith, L. C., Friedman, S., & Nevid, J. (1999). Clinical and sociocultural differences in African American and European American patients with panic disorder and agoraphobia. *Journal of Nervous and Mental Disease, 187,* 549–560.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for state-trait anxiety inventory*. Palo Alto: Consulting Psychologists Press.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811.

Stober, J. (1998). Reliability and validty of two widely-used worry questionnaires: Self-report and self-peer convergence. *Personality and Individual Differences, 24,* 887–890.

Tallis, F., Eysenck, M., & Mathews, A. (1992). A questionnaire for the measurement of nonpathological worry. *Personality and Individual Differences, 13*(2), 161–168.

Taylor, S., & Cox, B. J. (1998). An expanded Anxiety Sensitivity Index: Evidence for a hierarchic structure in a clinical sample. *Journal of Anxiety Disorders, 12,* 463–483.

Tulbure, B. T., Szentagotai, A., Dobrean, A., & David, D. (2012). Evidenced based clinical assessment of child and adolescent social phobia: A critical review of rating scales. *Child Psychiatry of Human Development, 43,* 795–820.

Turner, S. M., Beidel, D. C., Dancu, C. V., & Stanley, M. A. (1989). An empirically derived inventory to measure social fears and anxiety: The Social Phobia and Anxiety Inventory. *Psychological Assessment, 1,* 35–40.

Vittengl, J. R., & Holt, C. S. (1998). Positive and negative affect in social interaction as a function of partner familiarity, quality of communication, and social anxiety. *Journal of Social and Clinical Psychology, 17,* 196–208.

Warheit, G., Holzer, C., & Arey, S. (1975). Race and mental illness: An epidemiologic update. *Journal of Health and Social Behavior, 16,* 243–256.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology, 114*(4), 522–536.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*(4), 448–457.

Watson, D., Clark, L. A., & Carey, G. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Weems, C. F., Hammond-Laurence, K., Silverman, W. K., & Ginsburg, G. S. (1998). Testing the utility of the anxiety sensitivity construct in children and adolescents referred for anxiety disorders. *Journal of Clinical Child Psychology, 27,* 69–77.

Wolpe, J., & Lang, P. J. (1964). A Fear Survey Schedule for use in behavior therapy. *Behaviour Research and Therapy, 2,* 27–30.

Zinbarg, R. E., Mohlman, J., & Hong, N. N. (1999). Dimesnions of anxiety sensitivity. In S. Taylor (Ed.), *Anxiety sensitivity: Theory, research, and treatment of the fear of anxiety* (pp. 83–114). Mahwah: Erlbaum.

Assessment of Obsessive Compulsive Disorder with African Americans

10

Monnica T. Williams, Chad T. Wetterneck and Broderick Sawyer

Overview of OCD

Obsessive-compulsive disorder (OCD) is a highly disabling and distressing disorder, which has made it one of the leading causes of disability worldwide (Lopez and Murray 1998). Obsessions are intrusive, unwanted thoughts, images, or impulses that increase anxiety, whereas compulsions are repetitive behaviors or mental acts used to decrease anxiety (APA 2000). The disorder is equally common among men and women, causing significant and pervasive impairment in multiple domains, including home life, work, and relationships (Karno et al. 1988; Kessler et al. 2005; Ruscio et al. 2010). OCD-related costs have previously been estimated at US$ 8 billion annually in the USA (DuPont et al. 1995), and this figure would likely be higher today. OCD afflicts an estimated 1.6 % of the American population (Ruscio et al. 2010), and Black Americans experience OCD at rates equivalent to the general population (Himle et al. 2008; Ruscio et al. 2010; Zhang and Snowden 1999).

Quality of Life

A recent investigation of quality of life among African Americans found that those with OCD had the overall lowest life satisfaction when compared with individuals afflicted with other anxiety disorders and those with no disorder. Individuals with OCD also completed the fewest number of years in school, and this was significantly different from African Americans or European Americans with no disorder. There is also evidence of social impairments in African Americans with OCD, as the same study found that African Americans with OCD reported the least amount of closeness with friends when compared to those with other anxiety disorders (Bach et al. 2012).

Cultural Considerations for Diagnosing OCD

Comorbidity OCD is a highly comorbid disorder, with most sufferers also meeting criteria for another mood, anxiety, or substance use disorder (Ruscio et al. 2010). The National Survey of American Life (NSAL) found that a majority of Black Americans with OCD also met the criteria for at least one other psychiatric disorder, at 93.2 and 95.6 % in African Americans and Caribbean

---

M. T. Williams (✉) · B. Sawyer
Center for Mental Health Disparities, University of Louisville, Psychological & Brain Sciences, 2301 South Third Street, Louisville, KY 40292, USA
e-mail: m.williams@louisville.edu

B. Sawyer
e-mail: basawy01@louisville.edu

C. T. Wetterneck
Rogers Memorial Hospital, 34700 Valley Road, Oconomowoc, WI 53066, USA
e-mail: cwetterneck@rogershospital.org

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_10,

Blacks, respectively (Himle et al. 2008). A recent clinical study of African Americans with OCD found that 89.2 % had at least one other comorbid condition. Of these, the most prevalent were mood disorders (69.9 %), anxiety disorders (56.2 %), and substance abuse disorders (38.9 %). There was very low comorbidity with eating disorders; 5.5 % had a binge eating disorder and none met criteria for anorexia nervosa or bulimia nervosa (Williams et al. to appear). This is in stark contrast to findings with European Americans that tend to find high comorbidity rates between OCD and all eating disorders (e.g., Sallet et al. 2010).

Participants with hoarding compulsions as a primary symptom (43 %) were more likely to have a comorbid mood disorder than those without, were more likely to have a psychotic disorder and/or a substance abuse disorder, and were less likely to be diagnosed with another anxiety disorder (Williams, Davis, Thibodeau, and Bach, to appear).

Treatment Seeking and Barriers to Treatment Although 40 % of African Americans with OCD report their obsessional concerns to a doctor, very few receive treatment (Simmons et al. 2012), and African Americans are consistently underrepresented in OCD treatment clinics and research studies (Steever et al. 2012; Williams et al. 2008). Among all Americans, for those with severe OCD, 93 % receive some type of treatment (NCS-R; Ruscio et al. 2010); however, this is true for only 60 % for African Americans with severe OCD (NSAL; Himle et al. 2008), indicating a substantial health disparity. Even among those who are able to access mental health care, few African Americans with OCD receive specialized treatment, and only 20 % use a serotonin reuptake inhibitor medication (Himle et al. 2008). Both African American and European Americans report concerns about the cost of treatment, shame, stigma, and wanting to handle the problem on their own; however, certain barriers disproportionately affect African Americans, including not knowing where to find help and concerns about discrimination (Williams et al. 2012b).

Symptoms Dimensions It has been suggested one barrier to treatment could be a failure to identify OCD because of its heterogeneous presentation and possible cultural differences in symptom expression (Friedman et al. 2003; Sussman 2003). Variables related to OCD often differ cross-culturally (Williams et al. in press). In a clinical sample, Williams et al. (2012c) found that African Americans endorse obsessive-compulsive concerns in six major areas, comprising: Contamination and Washing, Hoarding, Sexual Obsessions and Reassurance, Aggression and Mental Compulsions, Symmetry and Perfectionism, and Doubt and Checking. These dimensions are similar to findings of studies in primarily White samples (i.e., Bloch et al. 2008; Williams et al. 2010). However, African Americans with OCD reported more contamination symptoms and are twice as likely to report excessive concerns with animals compared to European Americans with OCD. This indicates notable cultural differences, which is consistent with findings among nonclinical samples (Thomas et al. 2000; Williams and Turkheimer 2007).

Williams et al. (2012a) suggested that low socioeconomic status (SES) may be correlated to greater exposure and thereby concern about contaminants, and subsequently Williams et al. (2012c) found that lower income African Americans with OCD were significantly more concerned about contamination, lending support to that hypothesis. However, further investigation of possible cultural and historical factors that may impact symptom expression is needed, especially in light of the legacy of segregation, based on the notion that European Americans could become contaminated through close contact or sharing items with African Americans (e.g., Williams and Turkheimer 2007). Additionally, it has been hypothesized that European Americans may have developed stronger immune systems compared to African Americans through natural selection, as many of their ancestors survived the bubonic plague, possibly making them biologically less concerned about contaminants (Moalem et al. 2002).

African Americans express concerns about contamination from animals at twice the rate as

European Americans (Williams et al. 2012c). Williams and Turkheimer (2007) studied racial differences in OCD symptoms and found that a nonclinical sample of African Americans scored significantly higher on an animal attitude factor than European Americans (meaning they had greater concerns about animals), implicating cultural factors for this difference. It was hypothesized that perhaps the Western perspective of animals as pets is more socially acceptable among European Americans than other cultures that are more likely to regard animals as a source of food or vehicle for labor. Other cultural differences may relate to the historic practices such as the use of dogs as a means to hunt slaves or attack protesters during the Civil Rights era. This is consistent with recent work that suggests African Americans may experience greater phobias of animals (Chapman et al. 2008). As such, cultural differences are plausible contributing factors for increased animal sensitivity among those with OCD.

Fear of being misunderstood also seems to be more frequently endorsed by African Americans with OCD (Williams et al. 2012). An obsessive need to be perfectly understood could be a unique finding for African Americans related to fears of appearing unintelligent, resulting in *stereotype compensation*—an intentional effort to present one's self in a counter-stereotypical manner (Williams et al. 2008).

OCD Assessment Measures

Assessment of OCD is usually accomplished through a series clinical interviews and self-report measures. Here we discuss the most common assessment tools, their psychometric properties, and findings relevant specifically to African Americans.

Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) The most widely used OCD outcome measure is the Y-BOCS (Goodman et al. 1989a, b). The Y-BOCS is a semi-structured clinical interview that takes 45–60 min to complete, consisting of a comprehensive checklist of obsessions and compulsions and a 10-item severity scale. The checklist is most often administered before treatment and aids in treatment planning. The obsessions are listed in several categories including: aggressive (fear of harming others), contamination, sexual, hoarding/saving, religious, symmetry or exactness, somatic, and miscellaneous. The compulsions list is organized into categories including: cleaning/washing, checking, repeating, counting, ordering/arranging, hoarding/collecting, and miscellaneous. The factor structure of the checklist in African Americans is similar to findings in other ethnoracial groups (Bloch et al. 2008; Williams et al. 2012c).

The Y-BOCS severity scale rates the time occupied by obsessions and compulsions, how much they interfere with functioning, how much distress they cause, attempts to resist, and level of control. Items are rated on a 5-point scale ranging from 0 (no symptoms) to 4 (severe symptoms). The severity scale is usually administered pretreatment and may be administered again periodically throughout treatment and posttreatment. The total score is calculated by adding the ten items, yielding scores from 0–40.

The Y-BOCS severity scale shows good reliability ($\alpha = 0.88 - 0.91$) and validity in European American samples (Goodman et al. 1989a, b; McKay et al. 1995). Scores above 16 may be considered in the clinical range, and the mean for OCD patients is 21.9 ($SD = 8$). OCD severity can be further broken into severity ranges including: subclinical (0–7), mild (8–15), moderate (16–23), severe (24–31), and extreme (32–40) (Steketee 1999). Scores for healthy people without OCD tend to be very low ($M = 0.31$, $SD = 1.21$) (Simpson et al. 2006), although the measure was not intended for unafflicted individuals.

Very few studies have examined the Y-BOCS in African American samples. Two nonclinical studies examined the psychometric properties using the self-report version of the Y-BOCS in undergraduate and community participants (Garnaat and Norton 2010; Washington et al. 2008). Washington et al. (2008) concluded that a one-factor model was better than separating obsessions and compulsions. Garnaat and Norton (2010) compared four different ethnoracial groups and concluded the Y-BOCS may underestimate obsessions in African Americans compared to European American participants. One recent study examined the Y-BOCS in African Americans with OCD and found the overall scale demonstrated good internal consistency ($\alpha = 0.83$) and a significant positive relationship with other measures of psychopathology (Williams et al. 2013b). However, the best factor structure for the scale was a three-factor solution including severity of obsessions, resistance to obsession and compulsions, and severity of compulsions (Williams et al. 2013b). Based on the results of this study, the Y-BOCS appears to be a valid measure for African American populations and is recommended for use at this time.

Alternative forms of the Y-BOCS include a self-report version (Steketee et al. 1996) and a computerized version (Rosenfeld et al. 1992). The Y-BOCS has been translated into over 25 languages (Lam et al. 2005). The psychometric properties of the Y-BOCS self-report version have not yet been assessed in African Americans, so clinicians and researchers are advised against using these in this population as it is possible certain symptom dimensions (e.g., contamination), may result in inflated scores for non-pathological reasons.

There is also a YBOCS-II (Storch et al. 2010), which has been developed and validated in a sample of OCD afflicted individuals ($M = 20.48$, $SD = 11.11$). It shows strong reliability ($\alpha = 0.91$) and fair convergent validity with the OCI-R ($r = 0.69$), but it has not yet been validated in African Americans, and ethnoracial information was not specified in the validation sample of OCD patients. Although it is likely this measure is adequate for African Americans, use of a self-report version is likewise discouraged.

Children's Yale-Brown Obsessive-Compulsive Scale (CY-BOCS) The CY-BOCS (Scahill et al. 1997) includes symptom checklists and severity scales for both obsessions and compulsions that assess the presence of obsessions and compulsions in children and adolescents aged 8–18. The CY-BOCS 10-item clinician-rated severity scale has obsession ($M = 9.6$; $SD = 4.43$) and compulsion subscales ($M = 10.3$; $SD = 3.83$; 5 questions each), with items ranging from "0 = none", to "4 = extreme" and subscale scores ranging from 0 to 20, and total scores 0 to 40. The CY-BOCS yields strong psychometric properties for all 10 items ($\alpha = 0.87$; $M = 19.9$; $SD = 7.51$), for children and adolescents with OCD. The intraclass correlations for the CY-BOCS Total, Obsession, and Compulsion scores are 0.84, 0.91, and 0.66, suggesting high interrater agreement between the subscale and total scores. The ethnoracial composition of the validation sample was not reported.

Based on the study of the Y-BOCS, which found it to be an acceptable measure for African American adults (Williams et al. 2013b), the CY-BOCS is likely acceptable for African American youth. However, this remains to be established empirically.

Obsessive-Compulsive Inventory-Revised (OCI-R) The OCI-R is an 18-item self-report measure of distress from obsessions and compulsions (Foa et al. 2002). The total score ranges between 0 and 72. The questionnaire includes six subscales including: washing, checking, ordering, obsessing, hoarding, and neutralizing. The subscale scores range between 0 and 12. The OCI-R has shown good internal consistency ($\alpha = 0.81 - 0.93$), test-retest reliability ($r = 0.82 - 0.84$), and discriminant validity in the original validation sample of primarily European American patients (Foa et al. 1998, 2002). A clinical cutoff score of 21 differentiates OCD patients from nonpatients (Foa et al. 2002). The OCI-R has been translated into many languages, including Spanish (Fullana et al. 2005), Italian (Sica et al. 2009), Korean (Lim et al. 2008), German (Gönner et al. 2007, 2008), Icelandic (Smári et al. 2007), and French (Zermatten et al. 2006).

A recent study examined the psychometric properties of the OCI-R in a clinical and nonclinical sample of African Americans (Williams et al. 2013b). The originally proposed OCI-R six-factor structure exhibited good to excellent fit in the sample. Receiver operator characteristic (ROC) analyses indicated that a cut-off score of 36 exhibited the optimal balance of sensitivity and specificity—a score that is substantially greater than the cut-off score of 21 proposed in the original validation study (Foa et al. 2002). The higher score is consistent with previous studies of nonclinical samples of African Americans, which have produced greater scores on self-report measures for OCD (e.g., Williams et al. 2005, 2008). The OCI-R is suitable for use with African Americans provided that the higher cut-off score is employed.

The Obsessive Compulsive Inventory-Child Version (OCI-CV) The OCI-CV (Foa et al. 2010) is a 21-item self-report measure used to assess OC symptoms in children and adolescents from 7 to 17 years old. Items are scored on a 3-point Likert scale from "0 = never" to "2 = always," including questions such as: "I get upset if my stuff is not in the right order" or "If a bad thought comes into my head, I need to say certain things over and over."

The OCI-CV contains six subscales: doubting/checking ($M = 0.80$, $SD = 0.54$), obsessions ($M = 0.92$, $SD = 0.59$), hoarding ($M = 0.76$, $SD = 0.68$), washing ($M = 0.82$, $SD = 0.68$), ordering ($M = 0.82$, $SD = 0.58$), and neutralizing ($M = 0.69$, $SD = 0.65$); subscales can be summed to produce a total score ($M = 17.02$, $SD = 7.90$). Internal consistency was strong for the OCI-CV total and all subscale scores ($\alpha = \geq 0.81$), as was test-retest reliability over a period of one and a half weeks ($\alpha = \geq 0.68$). Additionally, the OCI-CV displays good convergent validity with the CY-BOCS total, obsession, and compulsion subscales ($r \geq 0.28$). OCI-CV scores appear to be sensitive to changes in symptom severity, similarly to the CY-BOCS. Although no studies have assessed the OCI-CV among nonclinical samples, mean scores among the Foa et al. (2010) clinical sample post-CBT treatment are as follows: doubting/checking ($M = 0.45$, $SD = 0.50$), obsessions ($M = 0.54$, $SD = 0.53$), hoarding ($M = 0.40$, $SD = 0.50$), washing ($M = 0.51$, $SD = 0.63$), ordering ($M = 0.54$, $SD = 0.49$), and neutralizing ($M = 0.37$, $SD = 0.50$), total score ($M = 10.16$, $SD = 7.58$).

The psychometric properties of the OCI-CV have not yet been examined in African American children and adolescents. Of the 109 children and adolescents in the validation study, only, 3.7 % were African American. Due to the scoring issues with the OCI-R (Williams et al. 2013a) and validity issues in similar self-report measures of OCD, the OCI-CV is not recommended for use in African Americans at this time.

Dimensional OCD Scale (DOCS) The DOCS (Abramowitz et al. 2010) is a self-report instrument consisting of 20 items. There are five items for each of the four symptom dimensions which include: (a) contamination, (b) responsibility for harm, injury, or bad luck, (c) unacceptable obsessional thoughts, and (d) symmetry, completeness, and exactness. An additional scale has been developed for the DOCS focused on sexually intrusive thoughts (DOCS-SI; Wetterneck et al. 2011). Within each symptom dimension, five items assess the following parameters of severity over the past month: (a) time occupied by obsessions and compulsions, (b) avoidance behavior, (c) associated distress, (d) functional interference, and (e) difficulty disregarding the obsessions and refraining from the compulsions. Each item is rated on a scale ranging from 0 (no symptoms) to 4 (extreme symptoms).

To accommodate the wide variety of OCD symptoms, the instructions for each set of five items include a brief description and several broad examples of the types of obsessions and compulsions observed in that dimension. Thus, the DOCS assess the severity of the patient's own symptoms rather than a predefined list of symptoms. The measure is designed for use in clinical and nonclinical populations and can be used by people aged 13 to adult.

The DOCS was validated within a primarily White adult sample of OCD patients ($M = 30.06$; $SD = 15.49$; 93 % European American), other anxiety disorder affected controls ($M = 16.75$; $SD = 13.14$; 92 % European American), and college undergraduates ($M = 11.93$; $SD = 9.87$; 77 % European American). The DOCS showed excellent reliability among all three populations (α's > 0.90), and good convergent validity when comparing subjects' total scores to OCI-R scores in all three samples (r's > 0.65). Additionally, the DOCS displayed sensitivity to treatment (exposure and ritual prevention), in similar ways to that of the Y-BOCS and OCI-R.

The DOCS has not yet been validated in a clinical sample of African Americans. However, there has been some attempt at validating the measure in a multicultural sample of college students. Wheaton and colleagues found that the DOCS overall and individual subscales demonstrated good to excellent internal consistency in African Americans (α = 0.84–0.89) (Wheaton et al. 2013). In addition, African Americans scored slightly higher, albeit not significantly, than Whites for contamination symptoms, but had no noticeable differences for harming, symmetry, or unacceptable thought symptom dimensions (Wheaton et al. 2013). The DOCS is not recommended for use in clinical samples of African Americans at this time. However, it may be useful among nonclinical populations.

Maudsley Obsessional Compulsive Inventory (MOC or MOCI) MOC or MOCI contains 30 dichotomously scored (true/false) items that assess obsessive-compulsive symptoms in the areas of contamination fears and washing behaviors, checking, and worries (Hodgson and Rachman 1977). The MOCI takes 5-min to complete and scores can range from 0 to 30. The means for OCD patients (Richter et al. 1994) and student samples (Dent and Salkovskis 1986) are 13.67 (*SD* = 6.01) and 6.32 (*SD* = 3.92), respectively. The reliability (Richter et al. 1994) and validity (Hodgson and Rachman 1977) are acceptable among Europeans and European Americans.

Thomas et al. (2000) found the MOCI lacked predictive validity in an African American undergraduate sample. African American students scored almost one standard deviation higher than European American students in total MOCI scores and African Americans scored significantly higher on the cleaning and checking subscales. Another study found similar differences in a combined student-community sample (N = 545; Williams and Turkheimer 2008). Although these studies did not use clinical samples, the findings raise serious concerns about the validity of the MOCI in assessing African Americans with OCD. African Americans without OCD may be appear to have clinically relevant symptoms.

Vancouver Obsessional Compulsive Inventory (VOCI) The VOCI (Thordarson et al. 2004) is a self-report, 55-item revision of the MOCI, which assesses obsessions, compulsions, avoidance behavior, and personality characteristics that often perpetuate symptoms of OCD. Participants are instructed to rate how true each statement reflects their thoughts and behaviors, with items ranging from 0 (not at all) to 4 (very much). Subscales include contamination, checking, obsessions, hoarding, just right, and indecisiveness. Total scores and subscales are calculated by summing the items. The VOCI was developed at the University of British Columbia, and distributed among four groups: OCD patients (*M* = 86.26; *SD* = 37.47), anxiety and depression controls (*M* = 49.61; *SD* = 43.28), community adults (*M* = 11.45; *SD* = 10.85), and students (*M* = 36.37; *SD* = 26.56). The ethnoracial composition of the sample was not described.

The VOCI displays strong evidence of validity and reliability for the total and subscale scores among European Americans; however, test-retest reliability is relatively low for scores in student samples, suggesting this measure may not screen low levels of OCD, and should be used primarily in clinical populations. Additionally, the VOCI's sensitivity to treatment has not yet been evaluated, and its psychometric properties have not yet been validated in African Americans.

Padua Inventories The original Padua Inventory contained 60 items about obsessions and compulsions on a 5-point rating scale in four main areas: contamination fears, checking, impaired control over mental activities, and worries about losing control over one's behaviors (Sanavio 1988). Two revised versions of the scale have been published including the 41-item Padua Inventory Revised (PI-R; Van Oppen et al. 1995) and the 39-item Washington State University Revision (PI-WSUR; Burns et al. 1996). The mean total score for individuals with OCD is 54.93 (*SD* = 16.72). The scale takes approximately 10-min to complete. Reliability and validity for the scale are good to excellent (Burns et al. 1996). The Padua Inventory is available in a variety of languages, including: English, Spanish (Mataix-Cols et al. 2002), Dutch (Van

Oppen 1992), Iranian (Goodarzi and Firoozabadi 2005), Japanese (Wakabayashi and Aobayashi 2007), and Korean (Min and Won 1999).

Studies of the Padua Inventory in African Americans found racial differences in mean item scores suggesting that the instrument does not measure the same traits in European American and African Americans (Williams et al. 2005). African Americans over reported cleaning and grooming behaviors compared to European Americans, which was attributed to different cultural backgrounds or beliefs regarding normal behaviors (grooming, washing, being cautious).

Another study examined the psychometric properties of the 10 contamination items, appearing both on the Padua Inventory and the PI-WSUR version in a large sample of community residents and undergraduate students (Williams and Turkheimer 2007). They found large racial differences on all items, and African Americans communicated more concern regarding contamination, cleanliness, and personal appearance. Williams et al. (2008) replicated the findings in a nationally representative sample, with African Americans scoring significantly higher than European Americans on the Padua contamination scale. The Padua Inventories have never been assessed in a clinical sample of African Americans, but based on findings with nonclinical samples, it seems this measure would be problematic for the assessment of African Americans.

Obsessive Belief Questionnaire (OBQ) The OBQ—Brief Version (OBQ-44; Steketee, 2005) is a 44-item self-report measure that assesses cognitive beliefs in OCD. The measure was revised using 44 items taken from the OBQ-87 to establish a smaller measure with three subscales (OCCWG 2001, 2003, 2005). Items are scored from 1–7 and calculated for a total score. The OBQ-44 consists of the following subscales: (1) Responsibility and threat estimation, (2) Perfectionism and intolerance for uncertainty, and (3) Importance and control of thoughts. In the initial study, Steketee (2005) administered the OBQ-44 to adult samples comprised of OCD patients ($M=174.3$; $SD=50.2$), anxious controls ($M=159.3$; $SD=53$), students ($M=131.3$; $SD=44.3$), and community dwellers ($M=96.0$; $SD=35.1$), finding strong reliability ($\alpha=0.95$) and validity when comparing PI-R subscales.

One study has examined the OBQ-44 in a clinical sample of African Americans with and without OCD. Williams et al. (2013b) found the total and three subscale scores of the OBQ-44 to have good to excellent internal consistency in African Americans with OCD ($\alpha=0.86$–0.94) and without ($\alpha=0.79$–0.96). In addition, the OBQ-44 demonstrated positive significant correlations to other measures of OCD severity including the Y-BOCS ($r=0.43$) and OCI-R ($r=0.64$). Thus, the OBQ-44 appears to be acceptable for use with African Americans, but more work needs to be done to validate the measure in this population.

Brown Assessment of Beliefs Scale (BABS) The BABS (Eisen et al. 1998) is a seven-item, semi-structured interview that assesses the degree of conviction and insight patients have concerning the beliefs underlying their obsessional thinking. Subscales include conviction ($M=2.5$; $SD=1.4$), perception of others' views of beliefs ($M=1.2$; $SD=1.5$), explanation of differing views ($M=1.9$; $SD=1.6$), fixity of ideas ($M=2.2$; $SD=1.4$), attempt to disprove beliefs ($M=2.1$; $SD=1.4$), insight ($M=1.3$; $SD=1.5$), and ideas/delusions of reference ($M=0.9$; $SD=1.5$). In the initial validation study, which included patients with OCD ($N=20$), body dysmorphic disorder ($N=20$), and mood disorders with psychotic features ($N=10$), overall interrater and test-retest reliability was excellent, as was internal consistency. One factor was found to account for 56% of the variance. Scores on the BABS also correlated with alternate measures of insight and were sensitive to changes in insight and treatment; however for symptom severity, scores were not related or identical to improvement.

Additionally, OCD patients may believe negative events will definitely occur if they do not perform specific compulsions. Considering this symptomatology, interviewers must be sure to phrase items that will properly assess these beliefs about these OCD-related compulsions, and their direct relationship to negative future events. This must be done by phrasing questions about future events with certainty, e.g., "How certain are you that you WILL (not might) contract HIV if you do not wash your hands repeatedly after touching a door knob?" (Eisen et al. 1998).

A recent study by Williams et al. (to appear) assessed the psychometric properties of the BABS in African Americans and found good internal consistency (α = 0.84), convergent validity with the Y-BOCS ($r = 0.47$), but a weak correlation with the OCI-R ($r = 0.17$) and OBQ-44 ($r = 0.24$). It is not yet clear if the BABS is a good measure for use with African Americans.

Clark-Beck Obsessive-Compulsive Inventory (CBOCI) The CBOCI (Clark et al. 2005) consists of 25-items that assess the frequency and severity of obsessive and compulsive symptoms. The measure has good internal consistency (α > 0.86) in clinical and nonclinical samples. In OCD patients (*M* = 42.14; *SD* = 16.07), it is correlated to other measures of OCD, such as the Y-BOCS ($r = 0.60$) and the PI-WSUR ($r = 0.65$), but it is also correlated to the BDI-II ($r = 0.57$), indicating questionable divergent validity. Furthermore, the measure was less reliable when administered to a student sample (*M* = 16.30; *SD* = 8.34), thus the authors note that caution must be exercised when using the CBOCI with a nonclinical sample.

As a symptom screener, the CBOCI is more similar to the OCI-R than other OC questionnaires, such as the PI-WSUR. Although it is difficult to assess the relative performance of these two measures without a direct comparison, the item content of the CBOCI differs, with 17 % of OCI-R items assessing obsessions compared with 56 % of the CBOCI items. The relatively equal attention given to both obsessions and compulsions leads the authors to conclude that the CBOCI is a more balanced measure of OCD symptoms.

The original validation sample was recruited from sites in the USA and Canada. The ethnoracial information provided by the authors separates participants into White and non-White groups, thus it is unclear how many, if any, African Americans were included. Additionally, the percentage of non-Whites is relatively small (6 %, $n = 29$), and no separate analyses were conducted for this group. Thus it is unknown whether the CBOCI is suitable for use with African Americans. Based on findings with other self-report symptom checklists, it is likely that African Americans will score higher and require a different cut-off score, as was seen with the OCI-R (Williams et al. to appear).

Florida Obsessive-Compulsive Inventory (FOCI) The FOCI (Storch et al. 2007) is a self-report measure that utilizes 25 items on two subscales: a Symptom Severity Scale and a symptom checklist. The Symptom Severity Scale (*M* = 12.76; *SD* = 3.52) uses items ranging from 0 (none) to 5 (extreme) to assess five aspects of severity (e.g., time spent on obsessive–compulsive symptoms), while the symptom checklist (*M* = 7.39; *SD* = 4.09) assesses the existence or lack thereof obsessions (10 items) and compulsions (10 items).

Within a sample comprised of only OCD affected individuals, the FOCI demonstrates strong convergent validity for the severity scale when compared to the Y-BOCS severity scale ($r = 0.61$), and the OCI-R ($r = 0.76$). In addition this scale shows good discriminant validity, strong internal consistency, and is sensitive to symptom decreases following CBT (Aldea et al. 2009a). In a nonclinical college student population (European American, 63.2 %; African American, 9.9 %), strong psychometrics were shown when comparing the FOCI severity scale with the OCI-R ($r = 0.58$; *M* = 3.9; *SD* = 3.3; Aldea et al. 2009b). However a similar measure of symptom severity (e.g., the Y-BOCS) was not measured in this study, so the validity of the symptom severity scale has not been fully assessed in nonclinical samples.

The FOCI would function well as a screener, as it is not only quick to administer, but displays the same strong psychometric properties across both OCD patient and nonclinical samples, correlating significantly with the Y–BOCS in each population. An advantage of the FOCI is its utilization of a subscale for measuring symptom severity. However, the FOCI may not be useful when collecting comprehensive information about individual OCD symptoms, as limited psychometric data are available, and has not yet been validated for African Americans. Because it may suffer from the same issues as similar measures, such as the OCI-R, Padua Inventories, and MOCI, it is not recommended for use with African American at this time.

Related OCD Assessment Measures

There are several additional measures that do not directly assess for OCD, but are useful for collecting data about OCD-related issues (e.g., barriers to treatment, family accommodation of symptoms, parental beliefs about childhood OCD). The most popular research and clinical instruments in this domain are described below.

Barriers to Treatment Questionnaire (BTQ) The BTQ (Goodwin et al. 2002; Marques et al. 2010) measures participants' perceived barriers to seeking OCD treatment, based on similar questionnaires from the broader literature. The BTQ assesses barriers to treatment in the following domains: logistic and financial, stigma, shame and discrimination barriers, and treatment perception and satisfaction barriers. OCD patients are asked if any of these possible barriers influenced them to not seek treatment. The BTQ has been used in diverse samples to identify OCD sufferers' barriers to treatment. Internal consistency for the BTQ was good for African Americans with OCD ($\alpha = 0.82$) and for a European American internet sample ($\alpha = 0.71$) (Marques et al. 2010; Williams et al. 2012b). Therefore, the BTQ is likely a good measure for use in African Americans.

Family Accommodation Scale (FAS) The FAS (Calvocoressi et al. 1995, 1999) is a 12-item semi-structured interview to assess the degree to which family members assist people with OCD in their rituals or avoidance behaviors. Family members' accommodation efforts are generally well-intentioned, however they often equate to more overall impairment and symptom severity (Storch et al. 2007). The FAS shows good internal consistency ($\alpha = 0.82$), as well as strong discriminant and convergent validity. Good validity has also been demonstrated when measuring family accommodation for pediatric OCD (Storch et al. 2007). In the adult validation studies conducted by Calvocoressi et al. (1995, 1999), no ethnoracial information was reported, while Storch et al. (2007) used a sample comprised of primarily European Americans (97 %).

Test-retest reliability and sensitivity to treatment have yet to be explored for this scale, and the FAS has not yet been validated in African Americans. Additionally, a recent exploratory study by Albert et al. (2010) demonstrated that family accommodation may be multidimensional, as PCA analyses indicated a three-factor structure (participation, modification, distress and consequences) for the FAS. Considering the lack of studies utilizing diverse samples, as well as the many unknown facets of family accommodation, the FAS is not recommended for use in African Americans at this time.

Parental Attitudes and Behaviors Scale (PABS) The PABS (Peris et al. 2008) is a 42-item parent-report scale designed to measure parent or caregiver responses to childhood OCD. It consists of three subscales: Accommodation, Hostility and Blame, And Empowerment. Parental attitudes and behaviors are measured utilizing a Likert-scale from Items 1 (not at all) to 5 (very often). In a sample of primarily European American families, the PABS demonstrates strong internal consistency for all three subscales, and good concurrent and predictive validity. The mean scores for the Accommodation, Hostility/Blame, and Empowerment subscales were 19.11 ($SD = 7.65$), 13.17 ($SD = 5.24$), and 15.86 ($SD = 4.28$), respectively. Although strong psychometric properties are evident, this scale has not been validated within African American populations, thus the PABS is not recommended for use among African Americans at this time.

Comprehensive Structured Interviews

Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) The SCID-I (First and Gibbon, 2004) is a semi-structured interview that diagnoses Axis I disorders in accordance with the DSM-IV. Questions assess past and present symptomatology, demographic, and other essential information (e.g., treatment history). When administering the SCID-I, a screening form consisting of 24 items may be given before the interview to assess symptomology for various Axis I disorders, if not, the interview in its entirety is administered, asking every question for all Axis I disorders. However with the screener, if patients select "yes" to either questions 8 (obsessions) and/or 9 (compulsions), the interviewer will go directly to that section to assess obsessive-compulsive symptom presence. During the interview, raters ask more about the obsessions (3 items) and compulsions (2 items), with responses ranging from 1 (absent or false) to 3 (threshold or true). If scoring a 3 for any item assessing symptom presence, the patient would then be prompted to describe the nature of their obsessions/compulsions. These follow-up questions include age of onset, how much stress the obsessions/compulsions are causing, whether or not the patient was on medication, etc. Then, clear directions to diagnose (and indicate severity) are presented on the assessment to the interviewer.

Diagnoses are determined by a varying set of criterion for each disorder, in which one would need a certain number of criteria to be diagnosed. The SCID-I also includes "leave out" instructions, enabling interviewers to omit questions about diagnoses if it is clear that a particular set of diagnostic criterion are not met. A shortened SCID-I is available, which assesses disorders which are more commonly seen in practice. However, the shortened version excludes certain disorders such as social phobia, eating disorders, and some specifiers such as 'with poor insight' for OCD (First and Gibbon, 2004). Studies find that the SCID-I displays fair reliability when assessing OCD (Zanarini and Frankenburg 2010; Lobbestael et al. 2010), but its psychometric properties have not been examined in African Americans.

The SCID is problematic for diagnosing OCD. One major issue is that it does not give enough examples of the different types of symptoms people with OCD might have, and thus participants may fail to recognize themselves in the descriptions provided. For example, the SCID says, "Now I would like to ask you if you have ever been bothered by thoughts that didn't make any sense and kept coming back to you even when you tried not to have them?" This statement could refer to just about anything, and it is critically dependent upon the interviewer to recognize what constitutes an "OCD thought" versus a worry related to generalized anxiety disorder, rumination associated with depression, a traumatic memory associated with posttraumatic stress disorder, or a normal intrusive thought to make an accurate diagnosis. Also, if the patient perceives that his/her OCD thoughts "make sense," they might not endorse the statement at all.

Even if the subject asks for clarification, there are only two examples given, (Thoughts like hurting someone even though you really didn't want to or being contaminated by germs or dirt?), but African Americans with OCD worry may worry about many different things (e.g., Williams et al. 2012c). Likewise, similar problems are evident in the question about compulsions. Therefore, the SCID is not recommended for the assessment of OCD in African Americans (Davis et al. 2013).

Anxiety Disorder Interview Schedule for the DSM-IV (ADIS-IV) The ADIS (Brown et al. 1994, 2001) is a semi-structured interview based on DSM-IV diagnostic criteria. The ADIS was developed to differentiate anxiety and other disorders that are commonly comorbid with anxiety disorders such as mood or substance abuse disorders (Summerfeldt and Anthony 2002). The ADIS has two versions, one that assesses current symptoms, and another that assesses current and past symptoms. Interviewers ask about personal demographics and specific symptoms for each disorder. Diagnoses are assigned a severity rating from 0–8, based on patient distress and impairment.

The ADIS displays strong internal consistency for both current ($\alpha = 0.85$) and lifetime ($\alpha = 0.75$) OCD symptom categories. However, threshold issues may arise when assessing key features of OCD, panic disorder, and specific phobia, which may lead to interrater disagreement as to whether there is sufficient evidence to assign principal diagnoses. There are no studies to date that have specifically assessed its validity in African

Americans with OCD. However, in a study assessing familial transmission of anxiety in African Americans, Chapman et al. (2012) indicated that the ADIS may be satisfactory when assessing anxiety in African American populations.

World Health Organization (WHO) Composite International Diagnostic Interview (CIDI) The CIDI is a structured interview developed for use with trained interviewers to assess mental disorders based on ICD-10 and DSM-IV diagnostic criteria. It can be utilized in epidemiological cross-cultural studies or for clinical and research purposes. The diagnostic portion of the CIDI is based on the World Health Organization's Composite International Diagnostic Interview (WHO CIDI 1990). In the 74-item OCD section of the CIDI, interviewers first ask yes/no questions to assess symptom presence. The first section assesses obsessions, and then the second assesses compulsions. For each section, if one or more symptom is present, the interviewer will ask a multitude of follow up questions (e.g., age of onset, levels of distress caused, interference with daily activities, levels of control over thoughts/behaviors, whether treatment was sought).

The CIDI shows strong reliability (Andrews and Peters 1998), but considering the methodology of large epidemiological studies, the CIDI-Auto has been developed and utilized in lieu of interviews that would require trained clinicians. This allows a larger amount of data collection, as the CIDI-Auto can be self-administered. Although useful in large studies, the reliability has been poor to fair when comparing CIDI-Auto and clinician rated diagnoses (Komiti et al. 2001).

The WHO version of the CIDI (WMH-CIDI) was used to assess lifetime and current Axis I disorders among African American and Caribbean Americans in the National Survey of American Life (NSAL) epidemiological study (Essau and Wittchen 1993; Jackson et al. 2004). The CIDI short-form version (CIDI-SF; Kessler et al.1998) was used for the OCD diagnostic module rather than the full WMH-CIDI OCD module (Himle et al. 2008). An OCD diagnosis from the CIDI-SF should be considered a probable diagnosis, since the section does not fully assess DSM-IV criteria. Specifically, persons with a positive CIDI-SF OCD are estimated to have an 84.2 % chance of having OCD according to the CIDI-SF OCD scoring guidelines. Thus, the use of the CIDI has not been fully explored in African Americans with OCD. However, the substance use version of the CIDI displayed strong reliability of scores between European and African Americans (Horton et al. 2000).

MINI International Neuropsychiatric Interview (MINI) The MINI (Sheehan et al. 1998) is a semi-structured clinical interview that provides information about the participant's psychiatric diagnoses (derived from DSM-IV and ICD-10), age of onset and comorbid conditions. Interviewees are presented with a few initial questions for each disorder (yes/no) about symptom presence, and queried with additional questions based on the presence/absence of symptoms assessed in the initial questions. In the OCD module of the MINI, there are two initial questions asked, one assessing obsessions and the other compulsions, within the past month. If either question is answered "yes," indicating symptom presence, patients are queried with a few follow up questions about the rationality of these behaviors, and a current diagnosis can be made. Unlike more extensive clinical interviews (e.g., CIDI, SCID-I), there are much fewer symptom presence/absence questions, as well as follow up items and current diagnosis criterion.

The MINI was intended to serve as a fast diagnostic tool for psychiatric disorders (15–20 min), although it may take up to an hour to administer. Additionally, unlike the CIDI and SCID-I, interviewers do not require extensive training to administer the MINI. The MINI also has other versions that can be used to assess children (MINI-Kid; Sheehan et al. 2010), act as fast screeners (MINI-Screen), or provide additional diagnostic information, similar to more extensive structured interviews (e.g., personality disorders; MINI-Plus). For OCD, the MINI displays good psychometric properties in European Americans, but has yet to be validated in African American samples.

At-a-glance summary table

Assessment name	*Disorder assessed*	*Recommendation(s) and/or relevant research findings*
Yale–Brown Obsessive Compulsive Scale—Checklist (Y-BOCS; Goodman et al. 1989a, b)	OCD symptoms in adults	The factor structure of the 60-item checklist was found to be similar in African American adults when compared to European American samples (Williams et al. 2012c)
Yale–Brown Obsessive Compulsive Scale—Severity Scale (Y-BOCS; Goodman et al. 1989)	OCD severity in adults	No ethnic/racial differences in the psychometric properties of the Y-BOCS severity scale were found in a non-clinical sample of African Americans (Washington et al. 2008). Symptom dimensions in African Americans are similar to other groups (Williams et al. 2012c)
Children's Yale–Brown Obsessive Compulsive Scale—checklist (CY-BOCS; Scahill et al. 1997)	OCD symptoms in children	No studies of this measure have been conducted with African American children
Children's Yale–Brown Obsessive Compulsive Scale—Severity Scale (CY-BOCS; Scahill et al. 1997)	OCD severity in children	No studies of this measure have been conducted with African American children
Maudsley Obsessional Compulsive Inventory (MOCI; Hodgson and Rachman 1977)	OCD severity in adults	Poor validity; African Americans reported higher total scores, with greater contamination and checking concerns. (Thomas et al. 2000)
Vancouver Obsessional Compulsive Inventory (VOCI;Thordarson et al. 2004)	OCD severity in adults	No studies of this measure have been done with African American adults with OCD. Not recommended for African Americans
Padua Inventory (PI, Sanavio 1988; PI-R Van Oppen et al. 1995; PI-WSUR, Burns et al. 1996)	OCD severity in adults	This measure has some problematic items and sub-scales when used with African Americans (Williams et al. 2005; Williams and Turkheimer 2007). African Americans report greater contamination anxiety and higher total scores. Also see Washington et al. (2008) and Williams et al. (2008)
Obsessive Compulsive Inventory, Revised (OCI-R, Foa et al. 2002)	OCD severity in adults	The OCI-R yields higher scores for African Americans on washing scale (Williams and Turkheimer 2007; Williams et al. 2008). Higher cut-off scores are needed for African Americans (Williams et al. to appear)
Obsessive Compulsive Inventory, Children's Version (OCI-CV, Foa et al. 2010)	OCD severity in children	This is a 21-item pediatric self-report measure of OCD severity, validated in youth aged 7–17. It includes subscales consisting of doubting/checking, obsessing, hoarding, washing, ordering, and neutralizing. Has not been validated in African American children
Dimensional OCD Scale (DOCS; Abramowitz et al. 2010)	OCD severity in adults	No studies of this measure have been conducted for African American adults with OCD
Obsessive Belief Questionnaire—Brief Version (OBQ-44; Steketee 2005)	Obsessional beliefs related to OCD severity in adults	The OBQ-44 demonstrates good to excellent internal consistency and significant correlations to other OCD severity measures in African Americans (Williams et al. 2013b)
Brown Assessment of Beliefs Scale (BABS; Eisen et al. 1998)	Conviction and insight in beliefs underlying obsessional thinking in adults	The BABS appears to have good internal consistency and significant correlations to other OCD severity measures in African Americans (Williams et al. 2013b)
Clark-Beck Obsessive-Compulsive Inventory (CBOCI; Clark et al. 2005)	OCD severity in adults	No studies of this measure with African American adults with OCD, although the CBOCI is similar to the OCI-R, and African Americans may require higher cutoff scores
Florida Obsessive-Compulsive Inventory (FOCI; Storch et al. 2007)	OCD severity in adults	No studies of this measure with African American adults with OCD

At-a-glance summary table

Related assessments	*Disorder assessed*	*Recommendation(s) and/or relevant research findings*
Barriers to Treatment Questionnaire (BTQ; Goodwin 2002; Marques et al. 2010)	Barriers to treatment of OCD	The BTQ is a short self-report measure that has been used in diverse samples, including African Americans (Williams et al. 2012b)
Family Accommodation Scale (FAS; Calvocoressi et al. 1999)	Family involvement in OCD behaviors	No studies of this measure with African American families caring for OCD afflicted individuals
Parental Attitudes and Behaviors Scale (PABS; Peris et al. 2008)	Parental involvement in OCD behaviors	No studies of this measure with African American parents caring for OCD afflicted children or adolescents

Summary

There are many measures to assess OCD, but only a few have been validated in African Americans. Self-report symptom checklists tend to produce inflated scores in African Americans, due to greater endorsement of contamination and/or checking items. Therefore, such measures should not be used with African Americans unless they have been specially validated for use in that group (e.g., the OCI-R; Williams et al. 2013a). Clinician administered measures, such as the Y-BOCS, appear to be more valid, but further studies are need to conducted to establish this conclusively.

References

Abramowitz, J. S., Deacon, B. J., Olatunji, B. O., Wheaton, M. G., Berman, N. C., Losardo, D., Timpano, K. R., McGrath, P. B., Riemann, B. C., Adams, T., Björgvinsson, T., Storch, E. A., & Hale, L. R. (2010). Assessment of obsessive-compulsive symptom dimensions: Development and evaluation of the dimensional obsessive-compulsive scale. *Psychological Assessment, 22,* 180–198. doi:10.1037/a0018260.

Albert, U., Bogetto, F., Maina, G., Saracco, P., Brunatto, C., & Mataix-Cols, D. (2010). Family accommodation in obsessive-compulsive disorder: Relation to symptom dimensions, clinical and family characteristics. *Psychiatry Research, 179,* 204–211. doi:10.1016/j.psychres.2009.06.008.

Aldea, M. A., Geffken, G. R., Jacob, M. L., Goodman, W. K., & Storch, E. A. (2009a). Further psychometric analysis of the Florida obsessive-compulsive inventory. *Journal of Anxiety Disorders, 23,* 124–129. doi:10.1016/j.janxdis.2008.05.001.

Aldea, M. A., Rahman, O., & Storch, E. A. (2009b). The psychometric properties of the Florida obsessive compulsive inventory: Examination in a non-clinical sample. *Individual Differences Research, 7,* 228–238.

American Psychiatric Association. (2000). *Diagnostic and statistical manual for mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.

Andrews, G., & Peters, L. (1998). The psychometric properties of the composite international diagnostic interview. *Social Psychiatry and Psychiatric Epidemiology, 33,* 80–88.

Bach, N., Simmons, A., Williams, M. T., Matusko, N., Abelson, J., & Himle, J. (Nov 2012). The quality of Life of African Americans living with anxiety disorders. Presented at the association of behavioral and cognitive therapies, National Harbor, MD.

Bloch, M. H., Landeros-Weisenberger, A., Rosario, M. C., Pittenge, C., & Leckman, J. F. (2008) Meta-analysis of the symptom structure of obsessive-compulsive disorder. *American Journal of Psychiatry, 165,* 1532–1542.

Brown, T. A., DiNardo, P.A., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV*. Albany: Graywind.

Brown, T.A., DiNardo, P.A., Lehman, C.L., & Campbell, L.A. (2001). Reliability of DSM-IV anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology, 111,* 49–58.

Burns, G., Keortge, S. G., Formea, G. M., & Sternberger, L. G. (1996). Revision of the Padua Inventory of obsessive compulsive disorder symptoms: Distinctions between worry, obsessions, and compulsions. *Behaviour Research and Therapy, 34,* 163–173. doi:10.1016/0005-7967(95)00035-6.

Calvocoressi, L., Lewis, B., Harris, M., Trufan, S.J., Goodman, W.K., McDougle, C.J., & Price, L.H. (1995). Family accommodation in obsessive-compulsive disorder. *American Journal of Psychiatry, 152,* 441–443.

Calvocoressi, L., Mazure, C., Kasl, S., Skolnick, J., Fisk, D.., Vegso, S.J., Van Noppen, B. L., & Price, L. H. (1999). Family accommodation of obsessive-compulsive symptoms: Instrument development and assessment of family behavior. *Journal of Nervous and Mental Disease, 187*(10), 636–642.

Chapman, L. K., Kertz, S. J., Zurlage, M., & Woodruff-Borden, J. (2008). A confirmatory factor analysis of specific phobia domains in African American and Caucasian American young adults. *Journal of Anxiety Disorders, 22,* 763–771.

Chapman, L. K., Petrie, J., Vines, L., & Durrett, E. (2012). The co-occurrence of anxiety disorders in African American parents and their children. *Journal of Anxiety Disorders, 26,* 65–70.

Clark, D., Beck, A., Antony, M., Swinson, R. & Steer, R. (2005). Screening for obsessive and compulsive symptoms: Validation of the Clark-Beck obsessive-compulsive inventory. *Psychological Assessment, 17*(2), 132–143. doi:10.1037/1040-3590.17.2.132.

Davis, D. M, Chasson, G., Combs, J., & Williams, M. T. (2013, November). The utility of the SCID in assessing obsessive-compulsive disorder in African Americans. *Poster presented at the Association of Behavioral and Cognitive Therapies*, Nashville, TN.

Dent, H. R., & Salkovskis, P. M. (1986). Clinical measures of depression, anxiety and obsessionality in non-clinical populations. *Behaviour Research and Therapy, 24,* 689–691. doi:10.1016/0005-7967(86)90066-5.

DuPont, R. L., Rice, D. P., Shiraki, S., & Rowland, C. R. (1995). Economic costs of obsessive-compulsive disorder. *Medical interface, 8*(4), 102–109.

Eisen, J. L., Phillips, K. A., Baer, L., Beer, D. A., Atala, K. D., & Rasmussen, S. A. (1998). The brown assessment of beliefs scale: Reliability and validity. *The American Journal of Psychiatry, 155,* 102–108.

Essau, C. A. & Wittchen, H.-U. (1993). An overview of the composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research, 3,* 79–85.

First, M.B., & Gibbon, M. (2004). The structured clinical interview for DSM-IV axis I disorders (SCID I) and the structural clinical interview for DSM-IV axis II disorders (SCID II). In: M. J. Hisenroth & D. L. Segal (Eds.), *Comprehensive handbook of psychological assessment: Vol 2. Personality assessment* (pp. 134–143). Hoboken: Wiley

Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E, & Amir, N. (1998). The validation of a new obsessive-compulsive disorder scale: The obsessive-compulsive inventory. *Psychological Assessment, 10*(3), 206–214.

Foa, E. B., Huppert, J. D., Leiberg, S., Langer, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The obsessive-compulsive inventory: Development and validation of a short version. *Psychological Assessment, 14*(4), 485–496.

Foa, E. B., Coles, M., Huppert, J. D., Pasupuleti, R. V., Franklin, M. E., & March, J. (2010) Development and validation of a child version of the obsessive compulsive inventory. *Behavior Therapy, 41,* 121–132.

Friedman, S., Smith, L. C., Halpern, B., Levine, C., Paradis, C., Viswanathan, R., & Ackerman, R. (2003). Obsessive-compulsive disorder in a multi-ethnic urban outpatient clinic: Initial presentation and treatment outcome with exposure and ritual prevention. *Behavior Therapy, 34,* 397–410. doi:10.1016/S0005-7894(03)80008-4.

Fullana, M. A., Tortella-Feliu, M., Caseras, X., Andión, Ó., Torrubia, R., & Mataix-Cols, D. (2005). Psychometric properties of the Spanish version of the obsessive-compulsive inventory-revised in a non-clinical sample. *Journal of Anxiety Disorders, 19,* 893–903. doi:10.1016/j.janxdis.2004.10.004.

Garnaat, S. L., & Norton, P. J. (2010). Factor structure and measurement invariance of the Yale-Brown Obsessive Compulsive Scale across four racial/ethnic groups. *Journal of Anxiety Disorders, 24*(7), 723–728.

Gönner, S., Leonhart, R., & Ecker, W. (2007). Das zwangsinventar OCI-R—die Deutsche Version des Obsessive-Compulsive Inventory-Revised: Ein kurzes selbstbeurteilungsinstrument zur mehrdimensionalen messung von zwangssymptomen. *Psychotherapie Psychosomatik Medizinische Psychologie, 57,* 395–404. doi:10.1055/s-2007-970894.

Gönner, S., Leonhart, R., & Ecker, W. (2008). The obsessive-compulsive inventory-revised (OCI-R): Validation of the German version in a sample of patients with OCD, anxiety disorders, and depressive disorders. *Journal of Anxiety Disorders, 22,* 734–749. doi:10.1016/j.janxdis.2007.07.007.

Goodarzi, M., & Firoozabadi, A. (2005). Reliability and validity of the Padua Inventory in an Iranian population. *Behaviour Research and Therapy, 43,* 43–54. doi:10.1016/j.brat.2003.11.004.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989a). The Yale-Brown Obsessive Compulsive Scale, I: Development, use, and reliability. *Arch Gen Psychiatry, 46,* 1006–1011.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., & Charney, D. S. (1989b). The Yale-Brown Obsessive Compulsive Scale, II: validity. *Arch Gen Psychiatry, 46,* 1012–1018.

Goodwin, R., Koenen, K. C., Hellman, F., Guardino, M., & Struening, E. (2002). Helpseeking and access to mental health treatment for obsessive-compulsive disorder. *Acta Psychiatrica Scandinavica, 106,* 143–149. doi:10.1034/j.1600-0447.2002.01221.x.

Himle, J. A., Muroff, J. R., Taylor, R. J., Baser, R. E., Abelson, J. M., Hanna, G. L., Abelson, J. L., & Jackson, J. S. (2008). Obsessive-compulsive disorder among African Americans and blacks of Caribbean descent: Results from the national survey of American life. *Depression and Anxiety, 25*(12) 993–1005.

Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behavior Research and Therapy, 15,* 389–395.

Horton, J., Compton, W., & Cottler, L. B. (2000). Reliability of substance use disorder diagnoses among African Americans and Caucasians. *Drug and Alcohol Dependence, 57,* 203–209.

Jackson, J. S., Torres, M., Caldwell, C. H., Neighbors, H. W., Nesse, R. M., Taylor, R. J., Trierweiler, S. J., & Williams, D. R. (2004). The National Survey of American Life: a study of racial, ethnic and cultural influences on mental disorders and mental heath. *International Journal of Methods Psychiatric Research, 13,* 196–207.

Karno, M., Golding, J. M., Sorenson, S. B., & Burnam, M. (1988). The epidemiology of obsessive-compulsive disorder in five US communities. *Archives of General Psychiatry, 45*(12), 1094–1099.

Kessler, R. C., Andrews, G., Mroczek, D., Ustun, T. B., Wittchen, H.-U. (1998). The world health organization composite international diagnostic interview short form (CIDI-SF). *International Journal of Methods in Psychiatric Research, 7*(4), 171–185.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593–602.

Komiti, A., A., Jackson, H. J., Judd, F. K., Cockram, A. M., Kyrios, M., Yeatman, R., Murray, G., Hordern, C., Wainwright, K., Allen, N., & Singh, B. (2001). A comparison of the Composite International Diagnostic Interview (CIDI-Auto) with clinical assessment in diagnosing mood and anxiety disorders. *Australian and New Zealand Journal of Psychiatry, 35*, 224–230.

Lam, R. W., Michalak, E. E., & Swinson, R. P. (2005). *Assessment scales in depression, mania, and anxiety*. London: Taylor & Francis.

Lim, J. S., Kim, S. J., Jeon, W. T., Cha, K. R., Park, J. H., & Kim, C. (2008). Reliability and validity of the Korean version of Obsessive-Compulsive Inventory—revised in a non-clinical sample. *Yonsei Medical Journal, 49,* 909–916.

Lobbestael, J., Leurgans, M., & Arntz, A. (2010). Inter-rater reliability of the structured clinical interview for DSM-IV axis I disorders (SCID I) and axis II disorders (SCID II). *Clinical Psychology and Psychotherapy, 18,* 75–79. doi:10.1002/cpp.693.

Lopez, A. D., & Murray, C. C. J. L. (1998). The global burden of disease, 1990–2020. *Nature Medicine, 4*(11), 1241–1243.

Marques, L., LeBlanc, N. J., Weingarden, H. M., Timpano, K. R., Jenike, M., & Wilhelm, S. (2010). Barriers to treatment and service utilization in an internet sample of individuals with obsessive-compulsive symptoms. *Depression and Anxiety, 27,* 470–475. doi:10.1002/da.20694.

Mataix-Cols, D., Sànchez-Turet, M., & Vallejo, J. (2002). A Spanish version of the Padua Inventory: Factor structure and psychometric properties. *Behavioural and Cognitive Psychotherapy, 30,* 25–36. doi:10.1017/S1352465802001042.

McKay, D., Danyko, S., Neziroglu, F., & Yaryura-Tobias, J. A. (1995). Factor structure of the Yale-Brown Obsessive-Compulsive Scale: A two dimensional measure. *Behaviour Research and Therapy, 33,* 865–869. doi:10.1016/0005-7967(95)00014-O.

Min, B., & Won, H. (1999). Reliability and validity of the Korean translations of Maudsley Obsessional-Compulsive Inventory and Padua Inventory. *Korean Journal of Clinical Psychology*, *18*, 163–182.

Moalem, S., Percy, M. E., Kruck, T. P., & Gelbart R. R. (2002). Epidemic pathogenic selection: an explanation for hereditary hemochromatosis? *Med Hypotheses, 59,* 325–329.

Obsessive Compulsive Cognitions Working Group. (2001). Development and initial validation of the obsessive beliefs questionnaire and the interpretation of intrusions inventory. *Behaviour Research and Therapy, 39*(8), 987–1006.

Obsessive Compulsive Cognitions Working Group. (2003). Psychometric validation of the Obsessive Beliefs Questionnaire and the Interpretation of Intrusions Inventory: Part I. *Behaviour Research and Therapy, 41*(8), 863–878.

Obsessive Compulsive Cognitions Working group. (2005). Psychometric validation of the Obsessive Beliefs Questionnaire and Interpretation of Intrusions Inventory—Part 2: Factor Analyses and Testing of a Brief Version.

Peris, T. S., Benazon, N., Langley, A., Roblek, T., & Piacentini, J. (2008). Parental attitudes, beliefs, and responses to childhood obsessive compulsive disorder: The Parental Attitudes and Behaviors Scale. *Child & Family Behavior Therapy, 30*(3), 199–214.

Richter, M. A., Cox, B. J., & Direnfeld, D. M. (1994). A comparison of three assessment instruments for obsessive-compulsive symptoms. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 143–147. doi:10.1016/0005-7916(94)90007-8.

Rosenfeld, R., Dar, R., Anderson, D., Kobak, K. A., & Greist, J. H. (1992). A computer-administered version of the Yale-Brown Obsessive-Compulsive Scale. *Psychological Assessment, 4,* 329–332. doi:10.1037/1040-3590.4.3.329.

Ruscio, A. M., Stein, D. J., Chiu, W. T., & Kessler, R. C., (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry, 15*(1), 53–63.

Sallet, P. C., de Alvarenga, P., Ferrão, Y., de Mathis, M., Torres, A. R., Marques, A., Hounie, A., Fossaluza, V., Conceicao do Rosario, M., Fontenelle, L., Petribu, K., & Fleitlich-Bilyk, B. (2010). Eating disorders in patients with obsessive-compulsive disorder: Prevalence and clinical correlates. *International Journal of Eating Disorders, 43*(4), 315–325.

Sanavio, E. (1988). Obsessions and compulsions: The Padua Inventory. *Behaviour Research and Therapy, 26,* 169–177. doi:10.1016/0005-7967(88)90116-7.

Scahill, L., Riddle, M. A., McSwiggin-Hardin, M., Ort, S. I., King, R. A., Goodman, W. K., Cicchetti, D., & Leckman, J. F. (1997). Children's Yale-Brown Obsessive Compulsive Scale: Reliability and validity. *Journal of American Academy of Child and Adolescent Psychiatry, 36,* 844–852.

Sheehan, D. V., Lescrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 58,* 22–33.

Sheehan, D. V., Sheehan, K. H., Shytle, R.D., Janavs, J., Bannon, Y., Rogers, J.E., Milo, K.M., Stock, S.L., & Wilkinson, B. (2010). Reliability and validity of the mini international neuropsychiatric interview for children and adolescents (MINI-KID). *J Clin Psychiatry, 71*(3), 313–26.

Sica, C., Ghisi, M., Altoè, G., Chiri, L., Franceschini, S., Coradeschi, D., & Melli, G. (2009). The Italian version of the Obsessive Compulsive Inventory: Its psychometric properties on community and clinical samples. *Journal of Anxiety Disorders, 23,* 204–211. doi:10.1016/j.janxdis.2008.07.001.

Simmons, A., Williams, M. T., Matusko, N., Abelson, J., Bach, N., & Himle, J. (Nov 2012). Help-seeking behaviors of African Americans with anxiety disorders. Presented at the association of behavioral and cognitive therapies, National Harbor, MD.

Simpson, H., Rosen, W., Huppert, J. D., Lin, S., Foa, E. B., & Liebowitz, M. R. (2006). Are there reliable neuropsychological deficits in obsessive-compulsive disorder? *Journal of Psychiatric Research, 40,* 247–257. doi:10.1016/j.jpsychires.2005.04.004.

Smári, J., Ólason, D. T., Eythórsdóttir, Á., & Frölunde, M. (2007). Psychometric properties of the Obsessive Compulsive Inventory-Revised among Icelandic college students. *Scandinavian Journal of Psychology, 48,* 127–133. doi:10.1111/j.1467-9450.2007.00574.x.

Steever, A. M., Leonard, R. C., Riemann, B. C., & Williams, M. T. (Nov 2012). Minority participation in a residential and intensive outpatient program for OCD. Presented at the association of behavioral and cognitive therapies, National Harbor, MD.

Steketee, G. (1999). *Overcoming obsessive-compulsive disorder. Best practice for therapy series*. Oakland: New Harbinger Publications.

Steketee, G. (2005). Psychometric validation of the obsessive belief questionnaire and interpretation of intrusions inventory—part 2: Factor analyses and testing of a brief version. *Behaviour Research and Therapy, 43*(11), 1527–1542.

Steketee, G., Frost, R., & Bogart, K. (1996). The Yale-Brown Obsessive Compulsive Scale: Interview versus self-report. *Behaviour Research and Therapy, 34,* 675–684.

Storch, E. A., Bagner, D., Merlo, L. J., et al. (2007). Florida Obsessive Compulsive Inventory: development, reliability, and validity. *Journal of Clinical Psychology, 63,* 851–859.

Storch, E. A., Geffken, G. R., Merlo, L. J., Jacob, M. L., Murphy, T. K., Goodman, W. K., Larson, M.J., Fernandez, M., & Grabill, K. (2007). Family accommodation in pediatric obsessive-compulsive disorder. *Journal of Clinical Child and Adolescent Psychology, 36,* 207–216.

Storch, E. A., Larson, M. J., Price L. H., Rasmussen, S. A., Murphy, T. A., & Goodman, W. K. (2010). Psychometric analysis of the Yale-Brown Obsessive-Compulsive Scale second edition symptom checklist. *Journal of Anxiety Disorders, 24*(6), 650–656.

Summerfeldt, L. J., & Antony, M. M. (2002). *Structured and semi-structured diagnostic interviews*. In M. M. Antony & D. H. Barlow (Eds.), Handbook of assessment, treatment planning, and outcome evaluation: Empirically supported strategies for psychological disorders (pp. 3–37). New York: Guilford.

Sussman, N. (2003). Obsessive-compulsive disorder: A commonly missed diagnosis in Primary Care. *Primary Psychiatry, 10*(12), 14.

Thomas, J., Turkheimer, E., & Oltmanns, T. F. (2000). Psychometric analysis of racial differences on the Maudsley Obsessional Compulsive Inventory. *Assessment,* 7(3), 247–258.

Thordarson, D.S., Radomsky, A.S., Rachman, S., Shafran, R., Sawchuk, C.N., & Hakstian, A. R. (2004). The Vancouver obsessional compulsive inventory (VOCI). *Behaviour Research and Therapy, 42,* 1289–1314

Van Oppen, P. (1992). Obsessions and compulsions: Dimensional structure, reliability, convergent and divergent validity of the Padua Inventory. *Behaviour Research and Therapy, 30,* 631–637. doi:10.1016/0005-7967(92)90008-5.

Van Oppen, P., Hoekstra, R. J., & Emmelkamp, P. G. (1995). The structure of obsessive-compulsive symptoms. *Behaviour Research and Therapy, 33,* 15–23. doi:10.1016/0005-7967(94)E0010-G.

Wakabayashi, A., & Aobayashi, T. (2007). Psychometric properties of the Padua Inventory in a sample of Japanese university students. Personality and Individual Differences*, 43,* 1113–1123. doi:10.1016/j.paid.2007.03.004

Washington, C. S., Norton, P.J., & Temple, S. (2008). Obsessive-compulsive symptoms and obsessive-compulsive disorder: A multiracial/ethnic analysis of a student population. *Journal of Nervous and Mental Disease, 196*(6), 456–461.

Wetterneck, C. T., Smith, A., Hart, J. M., & Burgess, A. (2011). Predictors of distress from sexually intrusive thoughts in a non-clinical population. *Journal of Cognitive Psychotherapy, 25,* 189–202.

Wheaton, M., Berman, N., Fabricant, L., & Abramowitz, J. (2013). Differences in obsessive compulsive symptoms and obsessive beliefs: A comparison between African Americans, Asian Americans, Latino Americans, and European Americans. *Cognitive Behavior Therapy, 42,* 9–20.

Williams, M. T., Chapman, L. K., Simms, J. V., & Tellawi, G. (in press). Cross-Cultural phenomenology of obsessive-compulsive disorder. In J. Abramowitz, D. McKay, & E. Storch (eds.), *Handbook of obsessive-compulsive disorder across the lifespan*. Wiley.

Williams, M. T., & Turkheimer, E. (2007). Identification and explanation of racial differences in contamination anxiety. *Behavior Research and Therapy, 45,* 3041–3050.

Williams, M. T. & Turkheimer, E. (2008). The effects of interviewer race on anxiety in African Americans.

In L. V. Sebeki (ed.), *Leading-edge health education issues*. New York: Nova Science Publishers.

Williams, M., Turkheimer, E., Schmidt, K. M., & Oltmanns, T. F. (2005). Ethnic identification biases responses to the Padua Inventory for obsessive-compulsive disorder. *Assessment, 12*(2), 174–185.

Williams, M. T., Turkheimer, E., Magee, E., & Guterbock, T. (2008). The effects of race and racial priming on self-report of contamination anxiety. *Personality and Individual Differences,* 44(3), 744–755.

Williams, M., Powers, M., Yun, Y. G., & Foa, E. (2010). Minority participation in randomized controlled trials for obsessive-compulsive disorder. *Journal of Anxiety Disorders,* 24(2), 171–177.

Williams, M. T., Abramowitz, J. S., & Olatunji, B. O. (2012a). The relationship between contamination cognitions, anxiety, and disgust in two ethnic groups. *Journal of Behavior Therapy and Experimental Psychiatry, 43,* 632–637. doi:10.1016/j.jbtep.2011.09.003.

Williams, M. T., Domanico, J., Marques, L., Leblanc, N., & Turkheimer, E. (2012b). Barriers to treatment of African Americans with OCD. *Journal of Anxiety Disorders—Special Issue on Ethnic Diversity, 26*(1), 555–563.

Williams, M. T., Elstein, J., Buckner, E., Abelson, J., & Himle, J. (2012c). Symptom dimensions in two samples of African Americans with obsessive-compulsive disorder. *Journal of Obsessive-Compulsive & Related Disorders, 1*(3), 145–152.

Williams, M. T., Davis, D., Thibodeau, M., & Bach, N. (2013a). Psychometric properties of the obsessive-compulsive inventory revised in African Americans with and without obsessive-compulsive disorder. *Journal of Obsessive-Compulsive & Related Disorders, 2*(4), 399–405.

Williams, M. T., Wetterneck, C. T., Thibodeau, M., & Duque, G. (2013b). Validation of the Yale-Brown Obsessive Compulsive Severity Scale in African Americans with obsessive-compulsive disorder. *Psychiatry Research 209*(2), 214–221. doi:10.1016/j.psychres.2013.04.007.

Williams, M. T., Brown, T., Bach, N., & Steever, A. (to appear). Psychiatric comorbidity in African Americans with obsessive-compulsive disorder and hoarding. Manuscript under review.

Williams, M., Davis, D., Thibodeau, M., & Bach, N. (in press). Psychometric properties of the obsessive-compulsive inventory revised in African Americans with and without obsessive-compulsive disorder. *Journal of Obsessive-Compulsive & Related Disorders*. doi:10.1016/j.jocrd.2013.07.003.

World Health Organization. (1990). *Composite international diagnostic interview*. Geneva: World Health Organization.

Zanarini M. C., & Frankenburg F. R. (2001). Attainment and maintenance of reliability of axis I and axis II disorders over the course of a longitudinal study. *Comprehensive Psychiatric, 42,* 369–374.

Zermatten, A., Van der Linden, M., Jermann, F., & Ceschi, G. (2006). Validation of the French version of the Obsessive-Compulsive Inventory-Revised. *European Review of Applied Psychology, 56,* 151–155.

Zhang, A. Y., & Snowden, L. R. (1999). Ethnic characteristics of mental disorders in five U.S. communities. *Cultural Diversity and Ethnic Minority Psychology, 5*(2), 134–146.

11

Assessment of Posttraumatic Stress Disorder with African Americans

Emily Malcoun, Monnica T. Williams
and Leyla Bahojb Nouri

Overview of PTSD

Posttraumatic stress disorder (PTSD) is a severe and chronic anxiety condition that may occur after experiencing or witness traumatic events. Among the general population, 60.7 % of men and 51.2 % of women are estimated to experience a traumatic event in their lifetime (Kessler et al. 1995), and a significant proportion of these develop PTSD. PTSD is characterized by reexperiencing, avoidance, and hyperarousal symptoms that occur over time and lead to significant disruption in one's life (American Psychiatric Association 2000). Although first acknowledged among combat veterans and rape victims, symptoms of PTSD are now recognized among those who survive auto accidents, sexual assaults, terrorist attacks, natural disasters, and even in first responders and medical professionals who care for those who have experienced a trauma (Friedman et al. 2011; Williams et al. 2010).

Epidemiology of PTSD in African Americans

It is important to note that the prevalence of PTSD differs by race and ethnicity. The National Comorbidity Survey-Replication (NCS-R) and National Survey of American Life (NSAL) found that, while non-Hispanic Whites are at higher risk for most anxiety disorders, African Americans show a 9.1 % prevalence rate for PTSD versus 6.8 % in non-Hispanic Whites (Himle et al. 2009). Among African Americans aged 55 and older, PTSD is the most prevalent disorder (Ford et al. 2007). Additionally, African Americans may become more impaired by PTSD than non-Hispanic Whites. African Americans showed significantly greater impairment in productivity and out-of-role domains, reflecting difficulty carrying out everyday activities (Himle et al. 2009).

Cultural Considerations for Diagnosing PTSD

Comorbidity

Among those with PTSD, the most common comorbid diagnoses are depressive disorders, substance use disorders, and other anxiety disorders (Brady et al. 2000). There is a substantial overlap between PTSD and symptoms of several other psychiatric disorders, which can contribute to missing the primary diagnosis of PTSD when trauma histories are not obtained (Brady et al.

M. T. Williams (✉) · L. B. Nouri
Center for Mental Health Disparities, Psychological & Brain Sciences, University of Louisville, 2301 South Third Street, Louisville, KY 40292, USA
e-mail: m.williams@louisville.edu

L. B. Nouri
e-mail: lvbaho01@louisville.edu

E. Malcoun
University of Pennsylvania, Department of Visual Studies, 3405 Woodland Walk 105 Jaffe History of Art Building, Philadelphia, PA 19104, USA
e-mail: emalcoun@mail.med.upenn.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_11,

2000). This problem may be particularly true for African Americans in mixed race counseling dyads, as the development of trust needed to disclose a trauma may be lacking in the absence of strong therapeutic alliance (i.e., Constantine 2007).

The comorbidity of PTSD with substance use disorders is complex because a substance use disorder may often develop as an attempt to self-medicate the distressing symptoms of PTSD. People with PTSD are at higher risk for nicotine dependence and drug abuse, although African Americans are not at a higher risk than European Americans (Breslau et al. 2006). Binge eating and obesity, both of which have been associated with PTSD, occur among African Americans at rates that are the same as or higher than those for European Americans (Franko 2007; Pagoto et al. 2012). Harrington et al. (2010) found a particularly strong link between PTSD and binge eating in African American women, related to the "strong Black woman" stereotype. Because the experience of PTSD can be dependent on cultural factors, it is important to conduct a proper, thorough assessment to accurately determine what is driving the symptoms.

Trauma Cognitions

PTSD changes the way people think about themselves, others, and the safety of the world. Research indicates that cognitions about these constructs may differ by race. For example. African Americans have lower expectations about the benevolence of the world in general (Zoellner et al. 1999). In a study of adults dually diagnosed with alcohol dependence and PTSD, it was found that African Americans had more negative thoughts about the safety of the world, fewer thoughts about self-blame, and similar levels of negative thoughts about the self (Williams et al. 2012). In addition, negative thoughts about one's self and self-blame about the trauma were connected to harmful patterns of drinking that impacted all areas of life broadly in African Americans, but this relationship was not evident in European Americans in the same study.

Treatment Seeking and Barriers to Treatment

The National Comorbidity Survey found that among respondents with PTSD there was an elevated risk of high school and college failure, becoming a teen parent, marital instability, and unemployment compared to people without PTSD (Brunello et al. 2001), although it is not clear if this is greater among African Americans. Given the greater overall difficulties faced by those with PTSD, accessing treatment is reasonably expected to be more challenging. Barriers to treatment among African Americans with PTSD include transportation difficulties, finances, family disapproval, and unfamiliarity with procedures for accessing treatment (Davis et al. 2008). Additionally, African Americans have higher feelings of stigma and more negative attitudes toward mental health treatment in general (Abdullah and Brown 2011), which is also a factor in reduced help-seeking.

One clinical study found that African Americans were significantly less likely to complete PTSD treatment compared with European Americans (Lester et al. 2010). African Americans had a 1.5 times greater likelihood of dropping out and a three times greater likelihood of never starting therapy. African American clients were more optimistic about the benefits of treatment prior to starting therapy. Furthermore, they were doing as well in treatment as their European American counterparts prior to dropout, even after accounting for demographic variables such as income and education, so lack of benefit from treatment does not explain this finding. The authors of the study attribute findings to the fact that many African Americans improved more quickly, stigma surrounding treatment, and lack of cultural sensitivity in the assessment and treatment process.

Racism and PTSD

One major factor in understanding PTSD in African Americans is the impact of racism on emotional and psychological well-being. The current criterion for a PTSD diagnosis implies that the event causing the distress must be negative and uncontrollable, and an individual's physical well-being is threatened (American Psychiatric Association [DSM-IV-TR] 2000). Although this description may address many forms of race-related trauma, it may not take into account ongoing low levels of racism that can lead to a general sense of distress and uncontrollability (Carter 2007). The latter experiences, though not physical in nature, attack the individual's identity and force the individual to relive traumas associated with their race or culture's history (Helms et al. 2011). Chouet al. (2012) found that perceived racial discrimination was associated with increased mental disorders in African, Hispanic, and Asian Americans, suggesting that perceived racism may in itself be a traumatic experience. Currently, the DSM recognizes racism as trauma only when an individual meets DSM criteria for PTSD in relation to a discrete racist event. This is problematic given that many African Americans experience cumulative experiences of racism as traumatic, with a discrete event acting as "the last straw" triggering trauma reactions (Carter 2007). Thus, current conceptualizations of trauma as a discrete event may be limiting for African Americans.

Moreover, existing PTSD measures aimed at identifying an index trauma fail to include racism among listed choice response options, leaving such events to be reported as "other" or fit into an existing category that may not fully capture the nature of the trauma (e.g. physical assault). This can be especially problematic since African Americans may be reluctant to report experiences of racism to European American therapists (Carter 2007; Constantine 2007), who comprise the majority of mental health clinicians (U.S. Department of Labor 2012). African American clients also may not link current PTSD symptoms to a single experience of racism if their symptoms relate to cumulative experiences of discrimination. For these reasons, therapists using existing PTSD measures with African Americans are encouraged to directly inquire about the client's experiences of racism when assessing trauma history. To ensure appropriate and accurate diagnoses, therapists must evaluate reactions to both discrete and cumulative experiences of racism.

If a mental health professional is not cognizant of race-based trauma, the symptoms and behavior may be misattributed to stereotypes about the client's race or culture (avoidance of strangers, aggressive reactivity; Helms et al. 2011). Because racism has not typically been considered a PTSD criterion A trauma, symptoms attributed to racist incidents may be downplayed or questioned, a mistake that only perpetuates the victim's anxieties (Carter 2007). Thus, clients may seek out mental healthcare to address race-based trauma, only to face further microaggressions (ambiguous racial slights) from their therapist (Sue et al. 2007).

Although it has been found that African Americans show more distress in response to racial harassment than they do to racial discrimination (Carter et al. 2005), knowledge of both kinds of experiences is essential to accurate assessment of PTSD in African Americans. Some forms of race-based traumas may include, but not be limited to witnessing ethnoviolence or discrimination of another, historical or personal memory of racism, institutional racism, microaggressions, and the constant threat of racial discrimination. The more subtle forms of racism mentioned may be predictable and commonplace, leading to constant vigilance, or "cultural paranoia," which may be considered a protective mechanism against the incidents (Carter 2007; Whaley 2001).

However subtle, these different forms of racism may result in victimization of an individual parallel to that induced by physical or threatened trauma. For example, Bryant-Davis and Ocampo (2005) addressed similar courses of psychopathology between rape victims and victims of racism. Similar to rape victims, race-related trauma victims may respond with dissociation, or shock, which can prevent them from responding to the incident in a functional manner. Victims may then feel shame and self-blame because they were unable to respond or defend themselves, which may lead to low self-concept and self-destructive behaviors (Bryant-Davis and Ocampo 2005). In the same investigation, a parallel was drawn between race-related trauma victims and victims of domestic violence. In both cases, survivors may feel shame over allowing themselves to be victimized. For instance, if those who experience a racist incident are told that if they work hard, dress a certain way, and get along well with others, they will not experience racism again. When these rules are followed, and racism still occurs, powerlessness, hyperarousal, and other symptoms associated with PTSD may set in or worsen (Bryant-Davis and Ocampo 2005).

PTSD Assessment Measures

Diagnosis and Assessment

PTSD is typically diagnosed in an individual who survived or witnessed actual or threatened death or serious injury, and experiences the following symptoms for a duration of greater than 1 month: (1) one or more reexperiencing symptoms (e.g., recurrent, distressing thoughts, images, flashbacks or dreams, intense emotional distress, or physical reactivity when reminded of the event); (2) three or more forms of trauma-related avoidance (e.g., trauma reminders, trauma-related thoughts and feelings, emotional numbing, inability to recall an important aspect of the trauma, markedly diminished interest in usual life activities, feeling detached or estranged from others, restricted emotional range, and a sense of a foreshortened future) and (3) at least two or more hyperarousal symptoms (e.g., sleep difficulties, irritability or outbursts of anger, hypervigilance, difficulty concentrating, and exaggerated startle response). Chronic PTSD is diagnosed in individuals with symptoms lasting greater than 3 months, while PTSD with delayed onset is diagnosed in those whose symptoms are triggered at least 6 months post trauma (American Psychiatric Association 2000).

Standardized clinical interviews and self-report questionnaires are used conjointly to assess (1) immediate reactions to the trauma (e.g., Immediate Stress Reaction Checklist, ISRC; Fein et al. 2001; Peritraumatic Dissociation Experiences Questionnaire, PDEQ-SR; Marmar et al. 1997); (2) evaluate the presence and severity of PTSD symptoms (e.g., Clinician-Administered PTSD Scale, CAPS; Blake et al. 1990; PTSD Symptom Scale Interview Version, PSS-I; Foa et al. 1993; PTSD Diagnostic Scale, PDS; Foa 1996); (3) allow the clinician to objectively assess trauma severity and differentiate symptoms such as remembering vs. reexperiencing (flashbacks) the event (e.g., PSS-I, Foa et al. 1993); and (4) evaluate previous trauma history (e.g., Standardized Trauma Interview, STI; Resick et al. 2008; Childhood Trauma Questionnaire, CTQ; Bernstein et al 1994). In addition to identifying trauma history and yielding a DSM diagnosis, the assessment process provides a therapeutic element by allowing the patient to begin relaying the details and impact of the trauma in a structured, gradual manner. Self-report measures are used in evidence-based treatment protocols, such as prolonged exposure therapy (Foa et al. 1997), to track changes in symptoms over the course of treatment and provide feedback on treatment progress and efficacy. In addition to its research utility, symptom tracking may encourage and provide feedback to patients in therapy asking them to face and discuss their traumatic memory.

This chapter summarizes measures designed solely for assessing PTSD in adults and children, focusing on their psychometric properties relevant to African Americans. Given the scope and focus of this chapter, comprehensive structured clinical interviews (e.g., Structured Clinician Interview for DSM-IV Axis I Disorders, SCID-I; First et al. 2002) and measures focusing on trauma-associated symptoms without evaluating PTSD (e.g., Dimension of Stressful Events Rating Scale, Fletcher 1996b; Posttraumatic Cognitions Inventory, Foa et al. 1999) are not included.

Clinician Administered Interviews

Clinician Administered PTSD Scale

The Clinician Administered PTSD Scale (CAPS; Blake et al. 1990) is a 30-item structured interview providing a current (past month) or lifetime diagnosis of PTSD, and may be used to assess symptoms during the past week. This measure evaluates 17 PTSD symptoms; symptoms related to social and occupational functioning; symptom changes since a previous CAPS administration, PTSD severity, response validity; and severity of five trauma-related symptoms (behavioral guilt, survivor guilt, lapses in consciousness, depersonalization, and de-realization). Criterion A is assessed using the Life Events Checklist (LEC), which identifies traumatic events experienced, with CAPS items asked in reference to up to 3 events. The CAPS can be administered by clinicians, researchers, and appropriately trained paraprofessionals. Standardized questions with probes are provided. The full interview takes 45–60 min to administer, but certain sections may be omitted (e.g., associated trauma symptoms). The measure was initially validated with 23 combat veterans, 19 of whom had previously been diagnosed with PTSD; however their race was unspecified. Internal consistency was acceptable for reexperiencing, numbing and avoidance, and hyperarousal (α = 0.77, 0.85, and 0.73, respectively; Blake et al. 1990).

The CAPS is regularly used to provide evidence of convergent validity and as a screening tool to detect PTSD across race (Goldmann et al 2011; Frueh and Kinder 1994; Kubany et al. 2000). In a study of traumatized, low-income African Americans, 21 % of the sample received a current PTSD diagnosis and 49 % received a lifetime PTSD diagnosis, with the highest rates in "severe" and "severe without amnesia" profiles, suggesting adequate validity (Nugent et al. 2012). In a mixed race sample of 126 male veterans (62.1 % European American, 37.9 % African American) previously diagnosed with PTSD, CAPS internal consistency was adequate (α = **0**.91), and no significant differences were found across race (Monnier et al. 2002). Similarly, Frueh et al. (2004) studied PTSD and comorbid disorders in a small sample of male veterans (57.1 % European Americans, 42.9 % African American) and found no significant differences in PTSD scores across race. Another study addressing VA service use by male veterans found similar results. In a sample with 35 % African Americans and 65 % European Americans, 12 % met the criteria for PTSD on the CAPS, with no significant racial difference in diagnosis or severity (Grubaugh et al. 2006). These results should be considered in light of study limitations, such as small sample size, short-term assessment period, and failure to control for trauma type (Monnier et al 2002).

While the CAPS is commonly used in studies with African Americans, none have adequately examined the validity and reliability of this measure in this population. Clinicians may choose to use this measure with their African American clients with this in mind. However, future research should focus on establishing the content validity of this measure before it can be confidently recommended for use in African Americans.

PTSD Symptom Scale-Interview Version

The PTSD Symptom Scale-Interview Version (PSS-I; Foa et al. 1993) evaluates the presence and severity of the 17 PTSD symptoms related to a single traumatic event. A 4-point scale is used to measure the severity of each symptom during the past 2 weeks. A PTSD diagnosis is made if an individual scores 1 or greater on one or more reexperiencing symptoms, 3 or more avoidance items, and 2 or more hyperarousal items. A total score between 10 and 20 is moderate, 21–28 is moderately severe, and above 28 is severe. Subscale scores for reexperiencing, avoidance and hyperarousal can also be summed.

The original validation sample was 71 % African American and had acceptable internal consistency for the total test (α = 0.85; Foa et al. 1993). Sample means were not reported. Powers et al. (2012) conducted a recent validation study of the PSS-I on a primarily African American sample (64.1 %) diagnosed with alcohol dependence and chronic PTSD. Internal consistency was satisfactory for the total severity (α = 0.90) and subscale scores (α = 0.74–0.85). PSS-I sample means were not reported in this study. Schumm et al. (2005) also documented excellent internal reliability (α = 0.94–0.95) in a study using the PSS-I with a community sample of mainly African American, low-income women who experienced sexual or physical childhood abuse.

The PSS-I has been used in several studies with African Americans as a diagnostic tool to determine eligibility for participation in clinical research studies (Zoellner et al. 1999). Since African Americans comprised of two validation samples, the PSS-I is acceptable for use with this population. An examination of its psychometric properties specific to African Americans would further substantiate its use with this group.

Self-Report Questionnaires

Childhood Trauma Questionnaire

The Childhood Trauma Questionnaire (CTQ; Bernstein et al. 1994) was originally a 70-item questionnaire assessing history of childhood abuse and neglect. The CTQ was developed from analyses of related material and is intended for adults and children in clinical settings. Participants were recruited from Mount Sinai Medical Center in New York City and the Veterans Affairs Medical Center in the Bronx. The CTQ consists of 4 scales (emotional/physical abuse, emotional neglect, sexual abuse, and physical neglect). Items assess the extent to which traumatic experiences were true when the respondent "was growing up," using a 5-point frequency rating scale (0 = never true, 4 = very often true).

The validation sample for the original version was predominantly African American (51.2 %) and male (85.3 %) with a history of substance abuse. Average raw total and scale scores were: total M = 110.8 (34.1); emotional/physical abuse M = 48.3 (18.6); emotional neglect M = 35.7, (12.5); sexual abuse M = 8.8 (5.3); and physical neglect M = 17.5 (6.2). Total and individual scales all showed acceptable internal consistency (α = 0.79, 0.95) and good test-retest reliability for total and scale scores was found when readministered to a randomly selected group of participants 2–6 months later (0.80–0.88).

Bernstein and colleagues later developed the 23-item CTQ brief version, in which a fifth scale was added by dividing physical and emotional abuse. Internal consistency was acceptable for the total score (α = 0.80) and scale scores (α = 0.79, 0.95). This screen was found to be generalizable across multiple samples (Bernstein et al. 2003). The CTQ brief version also demonstrated acceptable internal consistency in a study assessing history of child abuse in African American women residing in an urban setting (α = 0.79–0.95; Bradley et al. 2005). Based on the validation data, both versions of the CTQ are acceptable for use with African Americans. Further research into the psychometric properties with respect to African Americans and to subgroups within this ethnoracial population is needed to further substantiate its use among this group.

Davidson Trauma Scale

The Davidson Trauma Scale (DTS; Davidson et al. 1997) is a 17-item questionnaire assessing PTSD symptoms during the past week. Five point frequency (0 = not at all, 4 = every day) and severity scales (0 = not at all distressing, 4 = extremely distressing) evaluate PTSD in relation to an index trauma. Frequency and severity for overall PTSD and for each symptom cluster (intrusive reexperiencing, avoidance, hyperarousal) are totaled.

The DTS demonstrated good test-retest reliability ($r = 0.86$, $p < 0.001$), and excellent internal consistency for the overall test (α = 0.99), and for frequency and severity subscales, (α = 0.97, 0.98) in a validation sample of European American trauma survivors. Participants diagnosed with PTSD using the SCID had a mean raw score of 62.0 (SD 38.0), while those without a PTSD diagnosis had a mean of 15.5 (SD 13.8). Good internal consistency across scales was also documented in two separate studies of African American women who experienced a trauma (α = 0.88–0.90, Bliss et al. 2008; α = 0.79–0.90; Bradley et al. 2005). These studies support its use with African American women; however, no studies to date have examined its psychometric properties in African American men.

Distressing Event Questionnaire

The Distressing Event Questionnaire (DEQ; Kubany et al. 2000) is a 38-item self-report measure assessing DSM criteria A2-F for PTSD. Since the DEQ was designed to be used in conjunction with the Traumatic Life Events Questionnaire (Kubany et al. 2000), criterion A1 (exposure to actual or threatened death or serious injury) is assumed and therefore omitted. Remaining criteria are evaluated using a 5-point scale in 4 parts. Test developers found that the content validity of the

DEQ was representative and relevant to PTSD, however cutoff scores were inconsistent across groups (e.g., 26 for male combat veterans, 18 for female abuse survivors). These findings suggest that the psychometric properties for this measure need to be evaluated relevant to a specific population for which it is intended to be used. Given that African Americans were underrepresented overall, and that all African Americans (7%) in the validation sample were veterans (Kubany et al. 2000), the DEQ is not recommended for use with this group. Moreover, there are better alternatives that have been more extensively researched with African Americans.

Impact of Event Scale-Revised

The Impact of Event Scale-Revised (IES-R; Weiss and Marmar 1997) is a 22-item measure of perceived distress in relation to a specific traumatic event. Respondents report degree of distress by experienced in the last week by each "difficulty" listed using a 5-point scale (0=not at all, 4=extremely). The IES-R yields a total score (0–88) and subscale scores (intrusion, avoidance, and hyperarousal). While the IES-R is not used to diagnose PTSD, cutoff scores for a preliminary diagnosis of PTSD have been cited in the literature.

The developers examined the psychometric properties of the IES-R in emergency personnel who had responded to natural disasters in Los Angeles and San Diego and 1994 Northridge, CA earthquake survivors. High internal consistency was found across subscales in both groups ($\alpha=0.79$–0.92). Cut points and the ethnoracial composition of the validation sample were not specified. This measure displayed excellent internal consistency ($\alpha=0.96$) in a sample of 67 HIV-infected African American women (ages 18–45) from urban New Orleans (Kimmerling et al. 1999). Approximately 35% of the sample met criteria for PTSD using the IES-R, a number exceeding the general population (Kimmerling et al. 1999). Due to the size and medical status of the sample, these findings cannot be generalized. Further research is needed to substantiate validity of the IES-R in African Americans generally. Nonetheless, because the subscales are calculated, and the respondent can name and explain the triggering event; this measure may be of value in investigating symptoms associated with under-researched traumatic events, such as race-based incidents.

Los Angeles Symptom Checklist

The Los Angeles Symptom Checklist (LASC; King et al. 1995) is a 43-item measure of PTSD and related symptoms in adolescents and adults. The LASC does not focus on a specific trauma, rather it generally inquires about problems in the past month. The 17 DSM-IV PTSD symptoms are embedded among other items that assess general psychological distress. Items are rated using a 5-point scale (0=no problem, 4=extreme problem). A preliminary diagnosis of PTSD is given using DSM-IV criteria for symptoms rated 2 or higher. A severity score is computed by summing the ratings of the 17 PTSD symptoms. An index reflecting global assessment of distress and adjustment problems related to trauma exposure is obtained by summing all 43 items.

King et al. (1995) analyzed the psychometric properties of the LASC by aggregating data from 12 studies using multiple samples, including male Vietnam veterans, female survivors of intimate partner violence, and high-risk adolescents. Males and females were near-equally represented (52% female) in the pooled sample, but the ethnoracial composition was not reported. The LASC was highly internally consistent for aggregated sample and across groups varying on gender, age, and military versus civilian background for the PTSD (0.88–0.94) and full scale (0.94–0.95) indices. Sample means for the PTSD index were reported for samples and not for the pooled group and ranged from 12.3 (SD 9.0) for a sample of mixed gender adolescents (Burton et al. 1994) to 49.8 (SD 107.9) for male Vietnam Veterans (King et al. 1995). A study by Foy et al. (1997) found high internal consistency in a large sample of male and female adolescents for the 17-item PTSD index and the 43 item full scale ($\alpha=0.90$, 0.95). The sample was 25.8% African American and 48.5% Latino. Means for PTSD and full-scale indexes were 16.2 (SD 12.6) and 37.9 (SD 28.2) respectively. Normative information was provided for various adult and adolescent samples. The LASC has potential for use with various trauma groups; however its psychometric properties relevant to African Americans have not been investigated. Further research is needed to determine if it can accurately assess PTSD and associated features across race.

Modified PTSD Symptom Scale-Self Report Version

The 17-item Modified PTSD Symptom Scale-Self Report Version (MPSS-SR; Falsetti et al. 1993) is a modification of the PSS-I (Foa et al. 1993). Major changes to the previous version are: (1) items are not linked to a specific event, and (2) severity ratings are included for each item. Items are rated using 4-point frequency (0 = not at all, 3 = five or more times per week) and intensity scales (0 = not at all upsetting, 3 = extremely upsetting). Respondents rate items relevant to the past 2 weeks and link each symptom to a specific event when possible. The MPSS-SR can be used to make a preliminary determination of the diagnosis of PTSD using DSM-III-R criteria or a frequency, severity, or total score cutoff score, and can be interpreted as a continuous measure of PTSD symptom severity. Total and subscale (reexperiencing, avoidance, and hyperarousal) scores are computed by summing items.

Validation studies examined cutoff scores for treatment and community samples. Frequency cutoff scores were 23 for the treatment group and 15 for the community sample, severity cut points were 47 for the treatment group and 32 for the community sample, and total cutoff scores were 71 for the treatment group and 46 for the community sample. Internal consistency was excellent for both groups (treatment $\alpha = 0.96$, community $\alpha = 0.97$). The ethnoracial makeup of the groups was unspecified; however the PSS-I, on which the MPSS-SR is based, was validated on a primarily African American sample (70.7 %). Good internal consistency ($\alpha = 0.92$) was also found in a study of low-income African American trauma survivors (Gapen et al. 2011) and in a study assessing differences in PTSD symptoms across race of hurricane victims ($\alpha = 0.85$; Ai et al. 2011). While there is preliminary data suggesting that this measure is acceptable for use with African Americans, focused research on the psychometric properties of this measure specific to African Americans is needed.

Penn Inventory for Posttraumatic Stress Disorder

The Penn Inventory for Posttraumatic Stress Disorder (Penn Inventory; Hammarberg 1992) is a 26-item inventory assessing PTSD symptoms relevant to multiple traumas. This measure addresses PTSD symptoms that may not be specific to one event, and can be used with veterans, non-veterans, clinical populations, and nonclinical populations. Like the Beck Depression Inventory (BDI; Beck et al. 1961, 1988), respondents select the best fitting statement among four options per item. Each sentence measures the presence and severity or frequency of the related symptom and is scored 0–3. Scores range from 0–78, with a clinical cutoff score of 35 (Hammarberg 1992).

Excellent test-retest reliability ($r = 0.96$, $p < 0.001$) and internal consistency ($\alpha = 0.94$) were found in a mixed sample of veterans and non-veterans during and post treatment. Nearly half of the treatment (44 %) and posttreatment (48 %) groups were African American. In all, 20 % of veterans and 19 % of non-veterans were African American. Of those diagnosed with PTSD in the sample, 90 % were correctly identified using the Inventory, and a replication for cross-validation found 98 % sensitivity with the same groups (Hammarberg 1992). This measure appears to be acceptable for use with African Americans, and may be useful for assessing PTSD in relation to the cumulative impact of multiple race-based events, although further evaluation of the psychometric properties is needed.

Posttraumatic Diagnostic Scale

The PTSD diagnostic Scale (PDS; Foa 1996) is a 49-item measure used in clinical and research settings to evaluate PTSD in relation to an index trauma. The PDS assesses DSM-IV PTSD criteria A-F and requests a description of the index trauma. Symptoms during the last month are evaluated, though the assessment period may be modified for different purposes. The PDS consists of four sections. Part 1 is a trauma checklist that includes the option to endorse and specify some “other (unlisted)” traumatic event. In Part 2, respondents identify their most distressing traumatic event. Part 3 assesses the 17 PTSD symptoms using a 4-point scale (0 = not at all/only once, 3 = five or more times a week/almost always). Part 4 assesses symptom interference in major areas of life functioning. The PDS yields a total severity score (0–51) mainly reflecting the frequency of PTSD symptoms. A PDS Profile Report summarizes DSM-IV PTSD diagnostic status, totals the number of symptoms endorsed, and rates symptom severity and degree of impairment of functioning (Foa et al. 1997).

Approximately one third of the validation sample was African American (31 %). The mean total score was 33.6 (SD 10.0) for participants meeting criteria for PTSD and 12.5 (SD 10.5) for those who did not. Internal consistency was acceptable for the total (α = 0.92) and scale scores (α = 0.78–0.84). A study conducted by Hood and Carter (2008) also documented high internal consistency (α = 0.87) in a sample of 67 African American female abuse survivors. This tool may be useful for identifying race-based PTSD given that participants are given an opportunity to identify a traumatic event not included on the checklist. We recommend that assessors inquire about race-based trauma when completing the trauma checklist (Part 1), since African American clients may hesitate to volunteer incidents racism to White clinicians (Carter 2007; Constantine 2007). Otherwise, the PDS is likely an acceptable measure for use with African Americans, due to the large African Americans representation in the validation sample and subsequent work.

PTSD Checklist

The PTSD Checklist (PCL; Weathers et al. 1993) screens individuals for PTSD and measures symptom change during and after treatment. There are three versions of the PCL. The PCL-M (military) evaluates symptoms relevant to “stressful military experiences” in active Service Members and Veterans. The PCL-C (civilian) measures symptoms in relation to “stressful experiences” and can be used with any population in relation to multiple events, while the PCL-S (specific) assesses symptoms in relation to an identified trauma. A diagnosis can be derived in three ways: determining the presence of DSM-IV criteria, determining whether the severity score exceeds the clinical cutoff score, or combining these methods to confirm that an individual has the necessary pattern and sufficient severity symptoms required by the DSM-IV. A severity score is obtained by summing ratings across the 17 items corresponding to PTSD symptoms. The PCL is used in clinical and research settings.

The psychometric properties of the PCL were examined by Blanchard et al. (1996), who found that the instrument was a good brief screening measure for PTSD in a sample of 40 motor vehicle accident survivors and sexual assault victims. All items were significantly correlated with the CAPS. However, no information was provided about the ethnoracial composition of the sample.

The PCL showed good specificity and good sensitivity—correctly identifying 4 cases of PTSD, with 1 false positive, and 13 false negatives—in a sample of 51 trauma-exposed African Americans in the Detroit Neighborhood Health Study (Goldmann et al. 2011). More research is needed into the psychometric properties of all versions of this measure for African Americans in order to justify its use within this population. Clinicians may choose to use this measure with their African American clients with this in mind.

Screen for Posttraumatic Stress Symptoms

The Screen for Posttraumatic Stress Symptoms (SPTSS; Carlson 2001) consists of 17 items assessing PTSD relevant to multiple traumas. It was developed for screening PTSD symptoms in clinical and research settings. Respondents rate items using an 11-point frequency scale (0 = never, 10 = always). An overall score is computed by averaging ratings across the 17 items and subscales (reexperiencing, avoidance, and arousal). Since the items are not linked to a specific event, the SPTSS is appropriate for individuals with a history of multiple traumas or whose trauma history is unknown. An alternate version offers a response format assessing symptom frequency using a 5-point scale (0 = not at all, 4 = more than once a day).

A validation study was conducted with a sample of newly admitted psychiatric inpatients at a private hospital during a 3.5-year period. Participants were between the ages of 30 and 45, with 18 % identifying as African American and 79 % as European American. The mean overall score was 5.8 (SD 2.2). When an SPTSS cutoff score of 5.0 was established, sensitivity and specificity were adequate (0.85, 0.73), indicating good predictive validity (Carlson 2001). To date, the psychometric properties of the SPTSS with respect to African Americans have not been investigated, thus it is not recommended for this population. However, since this measure addresses multiple traumas, it has the potential for evaluating the cumulative impact of race-based incidents. Moreover, evidence suggesting that this is not an appropriate measure for use with African Americans is lacking.

Trauma Symptom Checklist-40

The Trauma Symptom Checklist-40 (TSC-40; Elliot and Briere 1992) is a 40-item measure of posttrauma reactions, used exclusively for research purposes. Respondents use a 4 point scale (0 = never, 3 = often) to rate how frequently each symptom was experienced in the last 8 weeks. The TSC-40 has 6 subscales (anxiety, depression, dissociation, sexual abuse trauma index, sexual problems, and sleep disturbances) and yields a total score ranging from 0–120. Validation sample data for this measure is currently unavailable.

The TSC-40 has been used in diverse samples, including educationally disadvantaged women enrolled in a residential treatment program. The sample included 50 African American women and 52 European American women and demonstrated good internal consistency ($\alpha = .93$; Ghee et al. 2010). Scores on the TSC-40 did not differ significantly between the African American and European American women ($M = 47.2$ [SD 23.0], $M = 48.9$ [SD 23.1], respectively; Ghee et al. 2010). A 5-factor model (negative mood, interpersonal problems, sleep disturbance, dissociative-like symptoms, and sexual problems) was determined to fit the African American and European American women in this sample better than the original 6-factor model, suggesting that the original factor structure may not be generalizable in diverse populations. These findings provide some support for using this measure with a subgroup of African American women. Further investigation of the psychometric properties of the TSC-40 specific to African Americans will further substantiate its use with other groups within this population.

Trauma Symptom Inventory

The Trauma Symptom Inventory (TSI; Briere 1996) is a 100-item global measure of trauma symptoms unrelated to a specific event. The TSI is intended for use in both clinical and research settings. The measure contains items corresponding to DSM-IV PTSD symptom criteria but does not specifically evaluate these criteria. Respondents rate frequency of each item in the past 6 months using a 4-point scale (0 = never to 3 = often). The TSI evaluates ten clinical scales corresponding to trauma related symptom domains (arousal, depression, anger/irritability, intrusive experiences, defensive avoidance, dissociation, sexual concerns, dysfunctional sexual behavior, impaired self-reference, and tension reduction behavior). Response validity using 3 scales is also provided (atypical responses, response level, and inconsistent response). Raw scores are converted to T scores for the clinical and validity scales using separate norms based on gender and age.

Briere et al. (1995) found acceptable internal consistency in a clinical validation sample consisting of 6.2 % African Americans. Alphas ranged from 0.74–0.90 across the 10 scales, with $\alpha = 0.87$. Means are available for each scale score, but total score means were not reported. Phelps et al. (2006) also documented good internal consistency for all scales ($\alpha = 0.87$–0.91) in a hospital-based study of 35 African American mothers of child abuse survivors. Additional studies psychometric properties of this measure are needed due to the small sample size and the potential bias of a convenience sample.

At-a-Glance Summary Table—Adult PTSD Measures

Assessment	Disorder or symptoms assessed	Recommendation(s) and/or relevant research findings
Childhood Trauma Questionnaire (CTQ; Bernstein et al. 1994)	Severity of abuse and neglect in childhood (retrospective)	Originally validated in a majority African American sample. Acceptable internal consistency in a sample of African American women for total (α=0.80) scale and subscales (α=0.79–0.95; Bradley et al. 2005)
Clinician-Administered PTSD Scale (CAPS; Blake et al. 1990)	PTSD diagnostic interview	While the CAPS has shown little to no difference in scores across race (Monnier et al. 2002; Frueh et al. 2004; Grubaugh et al. 2006), its psychometric properties specific to African Americans are unknown
Davidson Trauma Scale (DTS; Davidson et al. 1997)	PTSD symptoms	In a clinical sample of African American women, the severity and frequency scales of the DTS, as well as the total scale, showed strong internal consistency (α=0.79, 0.89, 0.90; Bradley et al. 2005)
Distressing Event Questionnaire (DEQ; Kubany et al. 2000)	PTSD symptoms	No studies of this measure have been conducted with African Americans
Impact of Events Scale-Revised (IES-R; Weiss & Marmar 1997)	Distress caused by traumatic events	Strong internal consistency in a small sample of HIV infected African American women (α=0.96; Kimmerling et al. 1999)
Los Angeles Symptom Checklist (LASC; King et al. 1995)	Distress caused by traumatic events	No studies of psychometric properties with respect to African Americans
Modified PTSD Symptom Scale (MPSS-SR; Falsetti et al. 1993)	PTSD symptoms in adults	Good internal consistency in low-income African American sample (α=0.92; Gapen et al. 2011)
Penn Inventory for Posttraumatic Stress Disorder (Penn Inventory; Hammarberg 1992)	PTSD symptoms	Appeared to accurately identify cases of PTSD in a sample of mixed race veterans (Hammarberg 1992)
Posttraumatic Diagnostic Scale (PDS; Foa et al. 1997)	Severity of PTSD symptoms	Acceptable internal consistency with African American female survivors of abuse (α=0.87; Hood & Carter 2008)
PTSD Checklist (PCL; Blanchard et al. 1996)	PTSD symptoms	Good internal consistency (α=0.93) in an African American sample and demonstrated 0.97 specificity and 0.24 sensitivity when administered along with the CAPS (Goldmann et al. 2011)
PTSD Severity Scale-Interview Version (PSS-I; Foa et al. 1993)	Severity of PTSD symptoms	Excellent internal consistency within a sample of low-income, predominantly African American women (α=0.94 and 0.95; Schumm et al. 2005)
Screen for Posttraumatic Stress Symptoms (SPTSS; Carlson 2001)	PTSD symptoms	No studies of psychometric properties with respect to African Americans

Assessment	Disorder or symptoms assessed	Recommendation(s) and/or relevant research findings
Trauma Symptom Checklist-40 (TSC-40; Elliot and Briere 1992)	Distress caused by traumatic events	Displayed excellent internal consistency in a sample of 102 educationally disadvantaged women, 50 identifying as African American (α=0.93; Ghee et al. 2010)
Trauma Symptom Inventory (TSI; Briere et al. 1995)	Exposure to trauma	Displayed good internal consistency across subscales with a sample of African American mothers of child abuse survivors (α=0.87–0.91; Phelps et al. 2006)

Child Measures

Children's Impact of Traumatic Events Scale-Revised

The Children's Impact of Traumatic Events Scale-Revised (CITES-2; Wolfe et al. 1991) is the most recent version of the CITES, a 78-item clinician-administered scale developed to assess the effects of sexual abuse on youths aged 8–16 years old. The CITES-R is comprised of 4 main factors and 11 subscales: (1) PTSD (intrusive thoughts, avoidance, hyperarousal, sexual anxiety), (2) Social Reactions (negative reactions from others, social support), (3) Abuse Attributions (self-blame and guilt, empowerment, personal vulnerability, dangerous world) and (4) Eroticism. Items are rated on a 3-point scale (0=not true, 1=somewhat true, 2=very true). In addition to the CITES-2, the 25-item CITES-Family Violence Form (CITES-FVF) assesses the effects of family violence on the child.

The internal consistency of the CITES scales were found to be moderately acceptable, with a mean alpha of 0.69 (α=0.56–0.79) in a validation sample of sexually abused children that were 75% female and 25% African American. Given that the psychometric properties of the CITES-R subscales are inconsistent and have not been investigated specific to African American children, it is not recommended for use with this population.

Child Posttraumatic Stress Reaction Index

The Child Posttraumatic Stress Reaction Index (CPTS-RI; Nader 1996) is a 20-item interviewer-administered scale for children aged 6–17 assessing PTSD symptoms as well as guilt, impulse control, somatic symptoms and regressive behaviors. Items are rated on a 5-point scale (0=none, 4=most of the time) reflecting symptom frequency, with total scores ranging from 0–80. Severity ranging from "doubtful" to "very severe" is based on the total scale score. Child and parent report versions are available and this measure is widely used in research and clinical practice. The CPTS-RI Revision 2 (also referred to as the PTSD Index for DSM-IV; Rodriguez et al. 2002) was developed to increase sensitivity for screening trauma exposure and criteria A1–A2, and provides more guidance for the assessor and child throughout the interview. There are several reaction index variations. Frequency response choices match to severity items.

The CPTS-RI demonstrated excellent internal consistency (α=0.97) in 110 African American children (Thompson and Massat 2005). Remaining psychometric properties with respect to Africans Americans still require investigation, thus it is not yet clear if this would be an acceptable measure to use with African Americans. Clinicians should use this measure with these limitations in mind.

Child PTSD Symptoms Scale

The Child PTSD Symptoms Scale (CPSS; Foa et al. 2001) is a 26-item version of the Posttraumatic Diagnostic Scale (PDS; Foa et al. 1997) for use with youth aged 8–18. The CPSS assesses PTSD diagnostic criteria in clinician administered (CPSS-I) and self-report (CPSS-SR) forms. It includes 2 event items, 17 symptom items, and 7 functional impairment items. Symptoms are rated on a 4-point frequency scale (0 = not at all, 3 = five or more times a week) and functional impairment items are scored dichotomously (present or absent). The CPSS yields a total symptom severity scale score (0–51) and a total severity-of-impairment score (0–7). Scores can also be calculated for each of the three PTSD symptom clusters.

Foa et al. (2001) documented high internal consistency in a sample of primarily European American children ($\alpha = 0.70$–0.89). Mean scores were 7.6 (SD 8.1) for the total scale, 1.9 (SD 2.7) for reexperiencing, 2.7 (SD 3.4) for avoidance, and 2.7 (SD 2.7) for arousal. Gillihan et al. (2013) found both versions of this measure to have variable internal consistency in a sample of female adolescent sexual abuse survivors, the majority of whom were African American (64 %). Alphas ranged from 0.58–0.83 for the CPSS-SR and 0.50–0.81 for the CPSS-I. There is data supporting the use this measure with African American female sexual assault survivors (Gillihan et al. 2013); however further research is needed to substantiate its use with African American male adolescents and in younger African American children.

Childhood PTSD Interview

The Childhood PTSD Interview (CPTSDI; Fletcher 1996a) is a 93-item semistructured interview that assesses PTSD and associated symptoms in relation to a single or multiple identified traumas in individuals aged 7–18. It includes a description of the traumatic event(s), PTSD symptom items, and associated symptom items. Child (CPTSDI-C) and parent (CPTSDI-P) versions are available. On the child version, all items are rated dichotomously (yes or no), while the parent version has several response formats including dichotomous selection and a 5–6-point Likert scale. The CPTSDI yields a categorical score of PTSD diagnosis as well as a continuous severity score obtained by adding all endorsed items. The questions are written at the third-grade reading level, but it has been used with younger children. A Parent Form assesses the same dimensions using language appropriate for adults in which a parent answers the questions with respect to the child's symptoms. The forms may also be used in a self-report format and either version can be administered by paraprofessionals. Validation sample data for this measure is currently unavailable, so its use cannot be recommended for African Americans at this time.

Children's PTSD Inventory

The Children's PTSD Inventory (CPTSDI; Saigh et al. 2000) is a clinician-administered scale for children aged 6–18 for assessing exposure to and symptoms related to a traumatic event. Examples of traumatic ("scary") experiences are described, and then the child is asked if he or she has ever experienced a scary event and, if so, if he or she felt upset when it happened and/or if the child felt he or she could do nothing to stop it from happening. If an event meets screening criteria then reexperiencing, avoidance/numbing, and arousal symptoms are assessed in reference to that event. Questions related to duration of symptoms are also asked.

The instrument yields dimensional and categorical scores to indicate severity and presence of a diagnosis of PTSD. In addition, scores on 5 subscales (Situational Reactivity, Reexperiencing, Avoidance and Numbing, Increased Arousal, and Significant Impairment) are provided. The CPTSDI can be used in either research or clinical settings. This measure is also available in Spanish. This instrument showed moderate to high alphas as (0.53–0.89) for the subscales and high internal consistency at the diagnostic level ($\alpha = 0.95$) in a sample of trauma exposed youth aged 7–18 (M = 13.8, SD = 2.9). Approximately 17 % of the validation sample was African American and the majority of participants were Hispanic American (65.7 %). Further research is needed to support use of this tool for assessing PTSD for African American youth.

Trauma Symptom Checklist for Children and Trauma Symptom Checklist for Young Children

The Trauma Checklist for Children (TSCC; Briere 1996) is a 54-item self-report scale designed for assessing trauma symptoms related to sexual abuse and other traumatic events in children aged 8–16. It consists of two validity scales (over- and underreporting of symptoms) and 6 clinical scales (Anxiety, Depression, Posttraumatic Stress, Sexual Concerns, Dissociation, and Anger). Items are rated on a 4-point frequency scale (0 = never, 3 = almost all the time). The TSCC is written at an 8-year-old reading level and has been normed for boys and girls. It is useful for evaluating of children who have experienced a range of traumatic events. The TSCC is available in two versions: the full 54-item test that includes 10 items tapping sexual symptoms and a 44-item alternate version (TSCC-A) that makes no reference to sexual issues. The instrument is suitable for individual or group administration, and it is suggested that a score of 60 or above is of clinical significance (Hunt et al. 2011; Briere 1996).

The Trauma Symptom Checklist for Young Children (TSCYC; Briere 2005) is a 90-item caretaker-report instrument that can be used to assess PTSD symptoms in children aged 3–12. It consists of 8 clinical scales (Anxiety, Depression, Anger/Aggression, Posttraumatic Stress—Intrusion, Posttraumatic Stress—Avoidance, Posttraumatic Stress—Arousal, Dissociation, and Sexual Concerns) as well as a summary PTSD scale (PTSD Total). A PTSD diagnostic worksheet provides a possible PTSD diagnosis in children aged 5 and older. It also contains two validity scales to assess caretaker over- and underreporting of the child's symptoms. Items are rated on a 4-point scale (1 = not at all, 4 = very often) in reference to the previous month. The TSCYC is normed separately for boys and girls within three separate age groups (3–4, 5–9, and 10–12). The TSCC and TSCYC can be administered and scored by paraprofessionals. Validation sample data, including sample means and cut points, is currently not available.

Hunt et al. (2011) administered the TSCC to a sample of African American children presenting at an urban mental health center from 2004 to 2007. The mean score was 48.1(SD 10.0), with 16 % placed in the clinically significant range. While these instruments have been used with African American youth, more research is needed to substantiate its use with this population.

The UCLA PTSD Index for DSM-IV

The UCLA PTSD Index for DSM-IV (UPID; Pynoos et al. 1998) is a revision of the CPTS-RI. This 48-item semistructured interview assesses exposure to 26 traumatic events and DSM-IV PTSD diagnostic criteria, and associated symptoms (guilt and fear of event's recurring). Severity scores of 38 or above are clinically significant. Steinberg et al. (2004) found this measure to be internally consistent (α = 0.90) across numerous studies in the USA and internationally.

Hunt et al. (2011) administered the UPID to a clinical sample of 257 African American children. The measure showed good convergent validity when correlated with other measures of PTSD (TSCC), although internal consistency was not reported in this sample. More research is needed to determine the validity of the UPID across race.

When Bad Things Happen Scale

The When Bad Things Happen Scale (WBTH; Fletcher 1996c) is 95-item self-report inventory. The questions in this scale exactly parallel the questions in the Childhood PTSD Interview and can be used as a complement to the interview. It includes 63 DSM-IV symptom items (3–6 per symptom) and 32 associated symptom items (2–5 questions per associated symptom).

Items are rated on a 3-point frequency scale (0 = never, 1 = some, and 2 = lots). The WBTH yields a categorical diagnosis of PTSD as well as a continuous severity score (a scoring system template is provided with the scale). It has a third-grade reading level and includes a parent report version. It is appropriate for use in either a research or clinical setting. However, there is no information available about the use of this measure in African American children.

At-a-Glance Summary Table—Child PTSD Measures

Assessment	Disorder assessed	Recommendation(s) and/or relevant research findings
Child Posttraumatic Stress Reaction Index (CPTS-RI; Nader 1996)	PTSD symptoms in children aged 6–17	High internal consistency (α = 0.97) in a sample of African American youth (Thompson Jr. and Massat 2005)
Child PTSD Symptom Scale (CPSS; Foa et al. 1997)	PTSD severity in children aged 8–18	High internal consistency found in mostly African American sample of female sexual assault survivors (α = 0.58–0.83 for CPSS-SR, α = 0.50–0.81 for CPSS-I; Gillihan et al. 2013)
Childhood PTSD Interview (Fletcher 1996a)	PTSD diagnostic interview for children aged 7–18	No studies of this measure have been conducted with African American children
Children's Impact of Traumatic Events Scale-Revised (CITES-2; Wolfe et al. 1991)	Effects of sexual abuse in children aged 8–16	No studies of this measure have been conducted with African American children
Children's PTSD Inventory (CPTSDI; Saigh et al. 2000)	PTSD diagnostic interview for children aged 6–18	No studies of this measure have been conducted with African American children
Trauma Checklist for Children (TSCC; Lanktree et al. 2008)	Trauma symptoms in children aged 8–16	M = 48.1 (10.0) in a clinical sample of African American youth (Hunt et al. 2011)
Trauma Symptom Checklist for Young Children (TSCYC; Lanktree et al. 2008)	PTSD symptoms in children aged 3–12	No studies of this measure have been conducted with African American children
The UCLA PTSD Index for DSM-IV (Pynoos et al. 1998)	Exposure to trauma in children aged 7–13 (Revision of Child PTSD Reaction Index)	Has been used with African American children with PTSD but not validated. (Hunt et al. 2011)
When Bad Things Happen Scale (WBTH; Fletcher 1996c)	PTSD symptoms in children aged 7–14 (at least a 3rd grade reading level)	No studies of this measure have been conducted with African American children

Utility of Existing Measures to Assess PTSD in African Americans

Of the 14 adult measures described in this chapter, 3 were validated in PTSD studies with a primarily African American sample. The psychometric properties of 8 have been investigated with respect to this population; however, typically only internal consistency was measured. Of the 9 child measures, none were validated using a primarily African American sample and psychometric properties of only 4 have been investigated relevant to this group. Thus, many PTSD measures have been used extensively in African American populations despite the lack of basic validation research. Clearly, many more studies investigating the reliability and validity of these measures are needed to substantiate the use of them with African American adults and children. Researchers should take deliberate steps to recruit adequate numbers of African Americans to ensure that measures can be validated in these and other ethnoracial populations (Williams et al. 2013).

Changes to the PTSD Criteria in DSM-5

Changes to the PTSD criteria for the DSM-5 have been made to ensure accurate diagnoses with new considerations. The first section involving the trauma experience has changed moderately, reflecting findings in clinical experience as well as empirical research. It is now within criteria if a person has learned about a traumatic event involving a close friend or family member, or if a person is repeatedly exposed to details about trauma (American Psychiatric Association 2013). The latter alteration was made to include those exposed in their occupational field, such as police officers, to remove ambiguities and strengthen the definition of a traumatic event (Friedman et al. 2011). A2 criterion, responding to the event with intense fear, helplessness, or horror has been removed. It was found that in many cases, such as soldiers trained in combat, emotional responses are only felt afterward, once removed from the traumatic setting (Friedman et al. 2011).

The most notable change to the criterion is from a three to a four-factor model. The DSM-5 factors are intrusion symptoms, persistent avoidance, alterations in cognition and mood, and hyperarousal and reactivity symptoms. These factors encompass the many facets of PTSD and were included based on empirical evidence (American Psychiatric Association 2013). These alterations will help distinguish PTSD from other conditions with similar symptomology, such as depression or acute stress disorder, while simultaneously expressing the condition on a spectrum (Friedman et al. 2011). Three new symptoms have been added—persistent distorted blame of self or others, persistent negative emotional state, and reckless or self-destructive behavior.

Two subtypes, acute and chronic, have been removed for lack of empirical evidence supporting the distinction (Friedman et al. 2011). In their place, a subtype has been proposed to include preschool children who may be experiencing and reacting to trauma differently than children over 6 years old. A second subtype addresses PTSD with Prominent Dissociative symptoms. This subtype includes those who may be putting more emphasis on avoidance behavior regarding the trauma, and so may not meet the original criteria. Recent rape victims may be a good example of this population (American Psychiatric Association 2013).

It is worth noting that many assessment instruments are administered by clinicians, which may be problematic when assessing PTSD cross-culturally. African Americans may express themselves nonverbally, and Parham (2002) suggests that nonverbal communication (e.g., body language and unexpressed emotion) should be measured when assessing African Americans, something that is rarely considered. Since language bias is a natural source of error in assessment, interviewers should be well-educated about cultural differences to accurately assess and diagnose the respondent.

As of this writing, no measures have been developed based on DSM-5 criteria, although new and revised measures will be developed in the coming months and years, with DSM-5 considerations in mind. Some of the DSM-5 changes open the door toward a broader understanding of PTSD and race-based trauma. New models are needed that incorporate this facet of the human experience, which are properly developed and evaluated with diverse populations in mind (Williams et al. 2014). The current state of scholarship in this area is unacceptable and calls for urgent action.

References

Abdullah, T., & Brown, T. L. (2011). Mental illness stigma and ethno cultural beliefs, values, and norms: An integrative review. *Clinical Psychology Review, 31*(6), 934–948.

Ai, A. L., Plummer, C., Kanno, H., Heo, G., Appel, H. B., Simon, C. E., & Spigner, C. (2011). Positive traits versus previous trauma: Racially different correlates with PTSD symptoms among Hurricane Katrina-Rita volunteers. *Journal of Community Psychology, 39*(4), 402–420.

American Psychiatric Association. (2000). *Diagnostic and statistical manual for mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental health disorders, fifth edition (DSM-5)*. Washington, DC: American Psychiatric Association.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77–100.

Bernstein, D. P., Fink, L., Handelsman, L., Foote, J., Lovejoy, M., Wenzel, K., Sapareto, E., & Ruggiero, J. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *Journal of Psychiatry, 151*(8), 1132–1136.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Taruna, A., Stokes, J., Handelsman, L., Medrana, M., Desmond, D., & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse and Neglect, 27*(2), 169–190.

Blake, D. D., Weathers, F., Nagy, L. M., Kaloupek, D. G., Klauminzer, G., Charney, D. S., & Keane, T. M. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS·1. *The Behavior Therapist, 13,* 187–188.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD checklist (PCL). *Behavioral Research & Therapy, 34,* 669–673.

Bliss, M. J., Ogley-Oliver, E., Jackson, E., Harp, S., & Kaslow, N. J. (2008). African American women's readiness to change abusive relationships. *Journal of Family Violence, 23*(3), 161–171.

Bradley, R., Schwartz, A. C., & Kaslow, N. J. (2005). Posttraumatic stress disorder symptoms among low-income, African American women with a history of intimate partner violence and suicidal behaviors: Self-esteem, social support, and religious coping. *Journal of Traumatic Stress, 18*(6), 685–696.

Brady, K. T., Killeen, T. K., Brewerton, T., & Lucerini, S. (2000). Comorbidity of psychiatric disorders and posttraumatic stress disorder. *Journal of Clinical Psychiatry, 61*(Suppl 7), 22–32.

Breslau, J., Aguilar-Gaxiola, S., Kendler, K. S., Su, M., Williams, D., & Kessler, R. C. (2006). Specifying race-ethnic differences in risk for psychiatric disorder in a USA national sample. *Psychological Medicine, 36*(1), 57–68.

Briere, J. (1996). *Trauma symptom checklist for children (TSCC)*. Odessa: Psychological Assessment Resources.

Briere, J. (2005). *Trauma symptom checklist for young children (TSCYC)*. Odessa: Psychological Assessment Resources.

Briere, J., Elliot, D. M., Harris, K., & Cotman, A. (1995). Trauma Symptom Inventory: Psychometrics and association with childhood and adult victimization in clinical samples. *Journal of Interpersonal Violence, 10*(4), 387–401.

Brunello, N., Davidson, J. R. T., Deahl, M., Kessler, R., Mendlewicz, J., Racagni, G., Shalev, A. Y., & Zohar, J. (2001). Posttraumatic stress disorder: Diagnosis and epidemiology, comorbidity and social Consequences, biology and treatment. *Neuropsychobiology, 43,* 150–162.

Bryant Davis, T., & Ocampo, C. (2005). Racist incident based trauma. *Counseling Psychologist, 33,* 479–500.

Burton, D., Foy, D. W., Bwanausi, C., Johnson, J., & Moore, L. (1994). The relationship between traumatic exposure, family dysfunction, and posttraumatic stress symptoms in juvenile offenders. *Journal of Traumatic Stress, 7,* 83–93.

Carlson, E. B. (2001). Psychometric study of a brief screen for PTSD: Assessing the impact of multiple traumatic events. *Assessment, 8,* 431–441.

Carter, R. T. (2007). Racism and psychological and emotional injury: Recognizing and assessing race-based traumatic stress. *The Counseling Psychologist, 35*(1), 13–105.

Carter, R. T., Forsyth, J. M., Mazzula, S. L., & Williams, B. (2005). Racial discrimination and race-based traumatic stress: An exploratory investigation. In R. T. Carter (Ed.), *Handbook of racial-cultural psychology and counseling, Vol 2: Training and practice* (pp. 447–476). Hoboken: Wiley.

Chou, T., Asnaani, A., & Hofmann, S. G. (2012). Perception of racial discrimination and psychopathology across three U.S. ethnic minority groups. *Cultural*

Diversity and Ethnic Minority Psychology, 18(1), 74–81.

Constantine, M. G. (2007). Racial microaggressions against African American clients in cross counseling relationships. *Journal of Counseling Psychology, 54*(1), 1–16.

Davidson, J. R. T., Book, S. W., Colket, J. T., Tupler, L. A., Roth, S., David, D., Hertzberg, M., Mellman, T., Beckham, J. C., Smith, R., Davison, R. M., Katz, R., & Feldman, M. (1997). Assessment of a new self-rating scale for post-traumatic stress disorder. *Psychological Medicine, 27,* 153–160.

Davis, R. G., Ressler, K. J., Schwartz, A. C., Stephens, K. J., & Bradley, R. G. (2008). Treatment barriers for low-income, urban African Americans with undiagnosed posttraumatic stress disorder. *Journal of Traumatic Stress, 21*(2), 218–222.

Elliott, D. M., & Briere, J. (1992). Sexual abuse trauma among professional women: Validating the Trauma Symptom Checklist- 40 (TSC-40). *Child Abuse and Neglect, 16,* 391–398.

Falsetti, S. A., Resnick, H. S., Resick, P. A., & Kilpatrick, D. (1993). The Modified PTSD Symptom Scale: A brief self-report measure of posttraumatic stress disorder. *The Behavioral Therapist, 16,* 161–162.

Fein, J. A., Kassam-Adams, N., Vu, T., & Datner, E. M. (2001). Emergency department evaluation of acute stress disorder symptoms in violently injured youths. *Annals of Emergency Medicine, 38*(4), 391–396.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002) *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P)*. New York: Biometrics Research, New York State Psychiatric Institute, November 2002.

Fletcher, K. (1996a). Psychometric review of the Childhood PTSD Interview. In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 87–89). Lutherville: Sidran.

Fletcher, K. (1996b). Psychometric review of dimensions of stressful events (DOSE) ratings scale. In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 144–151). Lutherville: Sidran.

Fletcher, K. (1996c). Psychometric review of the When Bad Things Happen Scale (WBTH). In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 435–437). Lutherville: Sidran.

Foa, E. (1996). *Posttraumatic diagnostic scale manual*. Minneapolis: National Computer Systems.

Foa, E., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress, 6,* 459–473.

Foa, E. B., Cashman, L., Jaycox, L., & Perry, K. (1997). The validation of a self-report measure of PTSD: The Posttraumatic Diagnostic Scale. *Psychological Assessment, 9,* 445–451.

Foa, E. B., Ehlers, A., Clark, D. M., Tolin, D. F., & Orsillo, S. M. (1999). The posttraumatic cognitions inventory (PTCI): Development and validation. *Psychological Assessment, 11,* 303–314.

Foa, E. B., Johnson, K. M., Feeny, N. C., & Treadwell, K. R. H. (2001). The Child PTSD Symptom Scale: A preliminary examination of its psychometric properties. *Journal of Clinical Child Psychology, 30,* 376–384.

Ford, B. C., Bullard, K. M., Taylor, R. J., Toler, A. K., Neighbors, H. W., Harold, W., & Jackson, J. S. (2007). Lifetime and 12-month prevalence of diagnostic and statistical manual of mental disorders, fourth edition, Disorders among older African Americans: Findings from the national survey of American life. *American Journal of Geriatric Psychiatry, 15*(8), 652–659.

Foy, D., Wood, J., King, D., King, L., & Resnick, H. (1997). Los Angeles symptom checklist: Psychometric evidence with an adolescent sample. *Assessment, 4*(4), 377–384.

Franko, D. L. (2007). Race, ethnicity, and eating disorders: Considerations for DSM-V. *International Journal of Eating Disorders, 40*(Suppl S3), S31–S34.

Friedman, M., Resick, P. A., Bryant, R. A., & Brewin, C. R. (2011) Considering PTSD for DSM-5. *Depression and Anxiety, 28*(9), 750–769.

Frueh, B. C., & Kinder, B. N. (1994). The susceptibility of the Rorschach inkblot test to malingering of combat-related PTSD. *Journal Of Personality Assessment, 62*(2), 280–298.

Frueh, B. C., Elhai, J. D., Monnier, J., Hamner, B. M., & Knapp, R. G. (2004). Symptom patterns and service use among African American and Caucasian veterans with combat-related PTSD. *Psychological Services, 1*(1), 22–30.

Gapen, M., Cross, D., Ortigo, K., Graham, A., Johnson, E., Evces, M., et al. (2011). Perceived neighborhood disorder, community cohesion, and PTSD symptoms among low-income African Americans in an urban health setting. *American Journal Of Orthopsychiatry, 81*(1), 31–37.

Ghee, A. C., Johnson, C. S., & Burlew, A. K. (2010). Measurement equivalence of the Trauma Symptom Checklist-40 for chemically dependent African American and Caucasian women: A preliminary study. *Journal of Aggression, Maltreatment & Trauma, 19,* 820–838.

Gillihan, S. J., Aderka, I. M., Conklin, P. H., Capaldi, S., & Foa, E. B. (2013). PTSD Symptom Scale: Psychometric properties in female adolescent sexual assault survivors. *Psychological Assessment, 25*(1), 23–31.

Goldmann, E., Aiello, A., Uddin, M., Delva, J., Koenen, K., Gant, L. M., & Galea, S. (2011). Pervasive exposure to violence and posttraumatic stress disorder in a predominantly African American urban community: The Detroit neighborhood health study. *Journal of. Traumatic Stress, 24*(6), 747–751.

Grubaugh, A. L., Frueh, C., Elhai, J. D., Monnier, J., Knapp, R. G., & Magruder, K. M. (2006). Racial differences in psychiatric symptom patterns and service use in VA primary care clinics. *Psychiatric Services, 57*(3), 410–413.

Hammarberg, M. (1992). Penn Inventory for posttraumatic stress disorder: Psychometric properties. *Psychological Assessment, 4,* 67–76.

Harrington, E. F., Crowther, J. H., & Shipher, J. C. (2010). Trauma, binge eating, and the "strong Black woman". *Journal of Consulting and Clinical Psychology, 78*(4), 469–479.

Helms, J. E., Nicholas, G., & Green, C. E. (2011). Racism and ethnoviolence as trauma: Enhancing Professional and research training. *Traumatology, 18,* 65–74.

Himle, J. A., Baser, R. E., Taylor, R. J., Campbell, R. D., & Jackson, J. S. (2009). Anxiety disorders among African Americans, Blacks of Caribbean descent, and non-Hispanic Whites in the United States. *Journal of Anxiety Disorders, 23*(5), 578–590.

Hood, S. K., & Carter, M. M. (2008). A preliminary examination of trauma history, locus of control, and PTSD symptom severity in African American women. *Journal of Black Psychology, 34*(2), 179–191.

Hunt, K. L., Martens, P. M., & Belcher, H. E. (2011). Risky business: Trauma exposure and rate of posttraumatic stress disorder in African American children and adolescents. *Journal Of Traumatic Stress, 24*(3), 365–369.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National comorbidity survey. *Archives of General Psychiatry, 52,* 1048–1060.

Kimmerling, R., Calhoun, K. S., Forehand, R., Armistead, L., Morse, E., Morse, P., Clark, R., & Clark, L. (1999). Traumatic stress in HIV-infected women. *AIDS Education and Prevention, 11*(4), 321–330.

King, L. A., King, D. W., Leskin, G., & Foy, D. W. (1995). The Los Angeles Symptom Checklist: A self-report measure of posttraumatic stress disorder. *Assessment, 2*(1), 1–17.

Kubany, E. S., Leisen, M. B., Kaplan, A. S., & Kelly, M. P. (2000). Validation of a brief measure of posttraumatic stress disorder: The distressing event questionnaire (DEQ). *Psychological Assessment, 12*(2), 197–209.

Lanktree, C. B., Gilbert, A. M., Briere, J., Taylor, N., Chen, K., Maida, C. A., & Saltzman, W. R. (2008). Multi-informant assessment of maltreated children: Convergent and discriminant validity of the TSCC and TSCYC. *Child Abuse & Neglect, 32*(6), 621–625.

Lester, K., Resick, P. A., Young-Xu, Y., & Artz, C. (2010). Impact of race on early treatment termination and outcomes in posttraumatic stress disorder treatment. *Journal of Consulting and Clinical Psychology, 78*(4), 480–489.

Marmar, C. R., Weiss, D. S., & Metzler, T. J. (1997). The peritraumatic dissociative experiences questionnaire. In J. Wilson & T. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 412–428). New York: Guilford.

Monnier, J., Elhai, J. D., Frueh, B. C., Sauvageot, J. A., & Magruder, K. M. (2002). Replication and expansion of findings related to racial differences in veterans with combat-related PTSD. *Depression and Anxiety, 16*(2), 64–70.

Nader, K. (1996). Psychometric review of childhood PTS reaction index (CPTS-RI). In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 83–86). Lutherville: Sidran

Nugent, N. R., Koenen, K. C., & Bradley, B. (2012). Heterogeneity of posttraumatic stress symptoms in a highly traumatized low income, urban, African American sample. *Journal of Psychiatric Research, 46*(12), 1576–1583.

Pagoto, S. L., Schneider, K. L., Bodenlos, J. S., Whited, M., Appelhans, B., Ma, Y., & Lemon, S. (2012). Association of post-traumatic stress disorder and obesity in a nationally representative sample. *Obesity, 20*(1), 200–205.

Parham, T. A. (2002). Counseling models for African Americans. In T. A. Parham (Ed.), *Counseling Persons of African Descent* (pp. 100–118). Thousand Oaks: Sage.

Phelps, L. F., Davies, W. H., McCart, M. R., Klein-Tasman, B. P., Melzer-Lange, M. D., & Heuermann, W. (2006). Concerns and coping of African-American mothers after youth assault requiring emergency medical treatment. *Journal of Pediatric Psychology, 31*(4), 388–396.

Powers, M., Gillihan, S. J., Rosenfield, D., Jerud, A., & Foa, E. B. (2012). Reliability and validity of the PDS and PSS-I among participants with PTSD and alcohol dependence. *Journal of Anxiety Disorders, 26,* 617–623.

Pynoos, R., Rodriguez, N., Steinberg, A., Stuber, M., & Frederick, C. (1998). *UCLA PTSD Index for DSM-IV. (Revision 1)*. Los Angeles: UCLA Trauma Psychiatry Program.

Resick, P. A., Galovski, T. E., Uhlmansiek, M., Scher, C. D., Clum, G., & Young-Xu, T. (2008). A randomized clinical trial to dismantle components of cognitive processing therapy for posttraumatic stress disorder in female victims of interpersonal violence. *Journal of Consulting and Clinical Psychology, 76,* 243–258.

Rodriguez, N., Steinberg, A., & Pyoonos, K. S. (2002). *The child posttraumatic stress reaction index, revision 2*. Los Angeles: UCLA Trauma Psychiatry Program.

Saigh, P., Yaski, A. E., Oberfield, R. A., Green, B. L., Halamandaris, P. V., Rubenstein, H., Nester, J., Resko, J., Hetz, B., & McHugh, M. (2000). The Children's PTSD Inventory: Development and reliability. *Journal of Traumatic Stress, 30,* 369–380.

Schumm, J. A., Stines, L. R., Hobfoll, S. E., & Jackson, A. P. (2005). The double-barreled burden of child abuse and current stressful circumstances on adult women: The kindling effect of early traumatic experiences. *Journal of Traumatic Stress, 18*(5), 467–476.

Steinberg, A. M., Brymer, M. J., Decker, K. B., & Pynoos, R. S. (2004). The University of California at Los Angeles post-traumatic stress disorder reaction index. *Current Psychiatry Reports, 6,* 96–100.

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., & Esquilin, M. (2007). Racial Microaggressions in everyday life: Implications for clinical practice. *American Psychologist, 62*(4), 271–286.

Thompson, T. r., & Massat, C. (2005) Experiences of violence, post-traumatic stress, academic achievement and behavior problems of urban African-American children. *Child & Adolescent Social Work Journal, 22*(5–6), 367–393.

U.S. Department of Labor. (Aug 2012). *Labor Force Characteristics by Race and Ethnicity*, *2011*, U.S. Bureau of Labor Statistics, Report 1036.

Weathers, F., Litz, B., Herman, D., Huska, J., & Keane, T. (1993). *The PTSD checklist: Reliability, validity, & diagnostic utility*. Annual Meeting of the International Society of Traumatic Stress Studies, San Antonio, TX.

Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale—Revised. In J. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 399–411). New York: Guilford.

Whaley, A. L. (2001). Cultural mistrust: An important psychological construct for diagnosis and treatment of African Americans. *Professional Psychology: Research and Practice, 32*(6), 555–562.

Williams, M., Cahill, S., & Foa, E. (2010). Psychotherapy for post-traumatic stress disorder. In D. Stein, E. Hollander, & B. Rothbaum (Eds.), *Textbook of anxiety disorders* (2nd ed.). Washington, DC: American Psychiatric.

Williams, M., Jayawickerme, N., Sposato, R., & Foa, E. B. (2012). Race-specific associations between trauma cognitions and symptoms of alcohol dependence in individuals with comorbid PTSD and alcohol dependence. *Addictive Behavior, 37*(1), 47–52.

Williams, M. T., Malcoun, E., Sawyer, B., Davis, D. M., Bahojb-Nouri, L. V., & Leavell Bruce, S. (2014). Cultural adaptations of prolonged exposure therapy for treatment and prevention of posttraumatic stress disorder in African Americans. *Journal of Behavioral Sciences, 4*(2), 102–124, doi:10.3390/bs4020102.

Williams, M. T., Tellawi, G., Wetterneck, C. T., & Chapman, L. K. (2013). Recruitment of ethnoracial minorities for mental health research. *The Behavior Therapist, 36*(6), 151–156.

Williams, M. T., Tellawi, G., Wetterneck, C. T., & Chapman, L. K. (in-press). Recruitment of ethnoracial minorities for mental health research. *The Behavior Therapist*.

Wolfe, V. V., Gentile, C., Michienzi, T., Sas, L., & Wolfe, D. A. (1991). The Children's Impact of Traumatic Events Scale: A measure of post-sexual abuse PTSD symptoms. *Behavioral Assessment, 13,* 359–383.

Zoellner, L. A., Feeny, N. C., Fitzgibbons, L. A., & Foa, E. B. (1999). Response of African American and Caucasian Women to Cognitive Behavioral Therapy for PTSD. *Behavior Therapy, 30*(4), 581–595.

Assessing Substance-Related Disorders in African Americans

12

Miesha Marzell, Brad Donohue and Travis Loughran

Introduction

Substance disorders are widespread, affecting approximately15 % of individuals in the general population (Hasin et al. 2007; Kessler et al. 2005). Caucasians generally evidence greater rates of substance abuse and dependence when compared to other racial and ethnic groups (Breslau et al. 2005; Breslau et al. 2006; Grant and Harford 1995; Kessler et al. 2005; Robins et al. 1984; Zhang and Snowden 1999). However, negative health outcomes related to substance abuse appear to disproportionately impact African Americans (Breslau et al. 2005). For instance, the mortality rate for cirrhosis of the liver due to alcohol abuse is higher in African Americans when compared to Caucasians (Caetano 2003). In addition, African Americans comprise approximately 14 % of the population of the USA, yet they represent nearly 21 % of admissions to public alcohol and drug treatment facilities (Substance Abuse and Mental Health Services Administration 2009). To address this apparent disparity, mental health service providers have increasingly considered ethnic culture in their assessment planning. Here it is important to note that in a comprehensive review of outcome studies for substance abuse, only 38 % of the controlled studies examined included African American participants, with half of the studies indicating 16 % or less African American representation (Strada et al. 2006). Similarly, psychometric development and dissemination of substance abuse assessment methodologies and interviewing protocols in African Americans has been slow to develop.

In response to this need, we report a practical, outlined approach to the assessment of substance disorders in African Americans. Our approach to substance abuse assessment in African Americans prescriptively follows evidence supported guidelines and culturally sensitive protocols. We conceptualize assessment as an ongoing process. General diagnostic considerations are presented to better understand symptoms of substance use disorders. We consider attitudes of African Americans toward substance abuse treatment, and potential barriers to seeking professional assistance. Because stigmatization and negative biases affect assessment in this population, we delineate strategies that may be utilized to assist African Americans in feeling more comfortable in the assessment process, such as involving community and family members in the assessment process. Evidence-supported methods of engaging

---

M. Marzell (✉)
Department of Community and Behavioral Health,
College of Public Health, The University of Iowa,
145 N. Riverside Drive, 100 College of Public
Health Bldg., Room N428, Iowa City,
Iowa 52242, USA
e-mail: miesha-marzell@uiowa.edu

B. Donohue · T. Loughran
Psychology Department, University of Nevada,
4505 S. Maryland Parkway, P.O Box 455030, Las Vegas,
NV 89154, USA
e-mail: bradley.donohue@unlv.edu

T. Loughran
e-mail: Travis.Loughran@unlv.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_12,

African Americans in the assessment process are emphasized, including those that have been formally evaluated to enhance session attendance. Of course, we review psychometrically validated measures commonly utilized in the assessment of substance abuse disorders, and potential issues that have been identified in their use with African Americans, such as clinical bias, validity.

Diagnostic Considerations

According to the Diagnostic and Statistical Manual of Mental Disorders (4th ed. text revision; DSM–IV) substance abuse involves a pattern of substance use leading to significant impairment or distress during the past 12 months (i.e., failure to fulfill major role obligations at work, school, or home; use in dangerous situations, legal problems, and continued use despite persistent or recurrent social or interpersonal problems; American Psychiatric Association 2000). Substance dependence is relatively more severe, involving a maladaptive pattern of substance use leading to clinically significant impairment. Substance dependence is reflected in the presence of several symptoms occurring at any time in the same *12-month period* (e.g., tolerance; withdrawal symptoms; using greater amounts than originally intended; persistent desire or unsuccessful efforts to reduce use; extended attempts to obtain or recover from substance; reduction everyday activities due to substance use; use despite persistent problems due to substance). The DSM-IV has been updated with a fifth edition. Essentially, substance abuse and dependence have been incorporated into a single diagnostic category (i.e., substance use disorder), with its severity being assessed along a continuum of endorsed symptoms (i.e., two to three symptoms = Mild disorder, four to five = Moderate disorder, six or more = Severe disorder; American Psychiatric Association 2013). Changes to the DSM-5 may be advantageous in the assessment of African Americans. For instance, recurrent legal problems, a problem for which African Americans have been found to be disproportionally affected in part due to biases in the legal system that are irrespective of crime, is no longer a criterion in the DSM-5, while objective criteria that do not appear to be influenced by race have been added (i.e., craving, strong desire, urge to use substance).

Although DSM-5 substance use disorder criteria incorporate both scientific evidence and clinical experience, inconsistent diagnostic reporting is a concern in ethnic minority samples (Minsky et al. 2003). Whereas, some investigators have determined that the DSM-IV criteria for substance dependence and abuse is reliable across ethnicities (Horton et al. 2000); other studies have found inconsistent results. For instance, the Substance Abuse Subtle Screening Inventory (SASSI; Miller 1985), an assessment tool designed to identify, substance disorders in high risk populations, has demonstrated relatively low predictive validity in African Americans as compared to Caucasians and Hispanics (Peters et al. 2000). The latter finding may be due to African Americans under-reporting substance use. Indeed, Perera-Diltz and Perry (2011) found African Americans are two to three times less likely to report using alcohol or drugs when compared to other groups. With regards to illicit drug use, studies have also demonstrated that there is a higher prevalence of underreporting among African Americans when compared to Caucasians and Hispanics (even when individuals are willing to submit to biological screening tests; Fendrich et al. 2004; Kim and Hill 2003; Richardson et al. 2003). Purported reasons for underreporting substance use include social desirability and cultural mistrust with those carrying out the diagnostic interviews (Johnson and Fendrich 2005). Further, African Americans often report greater negativity consequent to having reported substance use, as compared with other ethnic groups (Mensch and Kandel 1988). For instance, Rosenthal and Berven (1999) randomly assigned Caucasian graduate students to one of two groups. An African American client was presented to one group and a Caucasian client was presented to the other group. Other than ethnicity, the two client presentations were identical. Students provided clinical opinions about their respective client at two time points. At Time 1, minimal information was provided, and at time 2, comprehensive information was provided. The African American client was judged by the graduate students to have less potential for education and employment than the Caucasian client at both time periods. These results suggest assessors need to consider their propensity to derogatorily stereotype African Americans, and to consider evidence-based strategies designed to enhance trust and rapport in this population.

Substance Abuse Assessment Strategies

African Americans underutilize health services, including preventative health care (Fiscella et al. 2002). Particularly relevant to this chapter, mental healthcare in African Americans is underutilized due to its link to acute and chronic illnesses and death (Williams and Jackson 2005). Bolden and Wicks (2005) found that compared to other cultural groups, when African Americans do initiate psychiatric treatment it is more likely to be through emergency room admissions. African Americans also evidence significantly longer lengths of stay in treatment. This suggests that African Americans postpone mental health services until problems are relatively severe. Reasons have been put forth as to why African Americans underutilize therapeutic services, such as poverty, mistrust of the healthcare system, stigmatization, lack of knowledge regarding the therapeutic process, and the belief that African Americans do not need outside support systems to manage difficult life circumstances (Kessler et al. 1994; Sue and Sue 1990). To assist African American clients in their pursuit of mental health services, it is imperative that professionals address culture (Obasi and Leong 2009) and be culturally competent (Thompson et al. 2004).

Engagement During Initial Telephone Contact Our first contact with the African American client is usually over the telephone to assist in engagement, review program policies, and increase the likelihood of first session attendance (Donohue et al. 1999). We conduct this telephone call immediately after the referral is received because shorter wait periods improve treatment engagement (Claus and Kindleberger 2002). The telephone call includes a program introduction, statement as to who recommended the contact, and the professional's role at the agency. The professional solicits information regarding the referral, queries thoughts about the referral, and determines goals. Personal concerns are solicited, including potential problems with legal or social service systems. Lastly, personal goals and aspirations are celebrated. It is particularly important to focus on client goals, as some clinicians may hold negative biases regarding substance using clients (Taxman and Bouffard 2003; Treolar and Holt 2006). Avoiding problem focused assessment methods assists in normalizing therapy for African Americans who may otherwise think treatment is only appropriate for "crazy" people (Kranke et al. 2012). We replace therapeutic jargon and stigmatizing words with descriptive words that are positively focused (i.e., therapists, counselors, and psychologists are "professionals," therapy is "performance programming," sessions are "meetings", problem behaviors are "goal-worthy behaviors"). Orientating clients to therapy assists in reducing anxiety and improves engagement. It is also important to review treatment expectations. Milligan et al. (2004) found that African Americans who held treatment expectations that differed from their treatment plan were more likely to "drop out" of treatment than Caucasians who held the same belief. In scheduling meetings during the initial telephone contact, African American clients may feel session times are obligatory, and that they have little choice in determining alternative times to accommodate child care, non-flexible work schedules, and so on. Therefore, costs and scheduling options are reviewed.

Assuring Physical Layout of Clinic Celebrates Ethnic Culture Clinics need to be warm, inviting, and explicitly celebrate ethnic culture. It is important to assure art and pictures decorating walls represent a diversity of racial and ethnic groups. Staff should be trained to greet clients warmly by extending their hand, smiling, and offering snacks, beverages, and/or token gifts for children. Studies have shown that both clients and treatment staff perceive that the physical environment in which treatment is provided positively contributes to the completion of treatment objectives (Grosenick and Hatmaker 2000).

Solicitation of Family and Community Support Family, community, employer, and church have been shown to play vital roles in the assessment of substance abuse disorders within African American populations. Indeed, Brunswick et al. (1991) concluded that clinicians should consider acknowledgment of social ties and affiliations when assessing African Americans, especially in African American women. In contrast, they speculate the "revolving door" of treatment experienced by African American men may be a result of not addressing the important role of family, church, or community. Assessing the strengths and weaknesses of various social support systems makes it much easier to understand what social systems to enhance, and which individuals to involve in therapy sessions as supportive others.

Shortly after requisite consents are obtained in the first meeting, and sometimes during the initial telephone call, professionals assess appropriate support systems. These persons often include clergy, relatives, and close friends. Their role in the assessment process is to encourage and/or assist clients in attending therapy meetings and providing collateral information (which is particularly helpful in the assessment of undesired behaviors, e.g., substance abuse), encouraging effort, and providing praise for accomplishments. It may also be appropriate, to provide childcare services during intervention meetings at low rates or free of charge if possible.

Be Candid Reviewing the Purpose of Assessment In the assessment of substance use among African Americans it is important to assure psychometrically validated tests are utilized, and that clients are encouraged to express potential concerns with the referral process. Indeed, as compared with Caucasians, African American clients are less likely to seek treatment on their own and are more likely to be referred by a third-party (Jacobson et al. 2007). It is important for the assessor to provide a detailed rationale for testing, and explain the legal process as it relates to sharing the assessment results with others (i.e., release of information from client, court mandating, and duty to warn). Prior to administering each instrument, the professional explains what the respective test is purported to assess. In this way African American clients are less likely to feel anxious and/or defensive in responding candidly. It is equally important to answer questions in a straightforward, matter of fact manner, appreciating that it may be especially important to invite comments and questions as African Americans may be more cautious due to historical abuses of African Americans within the US health care system (Freimuth et al. 2001).

Assessment of Background Information Prior to administering psychometrically validated tests and measures, background information is assessed, including basic demographic information, educational and vocational experiences, hobbies, family constellation and relationships, medical and mental health history. Of note, African Americans disproportionately evidence diseases when compared with Caucasians. For example, African Americans are diagnosed with diabetes more than twice that of Caucasians, have one of the highest rates of hypertension in the world, and are more likely to die from heart disease as compared to Caucasians (Menke et al. 2013; Lloyd-Jones et al. 2010). Similarly, HIV infection is more prevalent in African Americans. In 2010, more than 40 % of all new HIV infection cases occurred in African Americans (Centers for Disease Control and Prevention (CDC) 2012). African American women between the ages of 25 and 34 are particularly affected, as HIV infection is the number one cause of death for this age group (Centers for Disease Control and Prevention 2012). Therefore, this demographic should be assessed for HIV and various risk factors associated with HIV (e.g., unsafe sex, intravenous drug use). It is also important to appreciate that many African Americans may be less likely than Caucasian clients to openly discuss medical concerns due to historical violations of their care within these settings (Shavers et al. 2002). Thus, it is especially important to indicate to African American clients how medical history may be associated with psychologically-based assessment. For instance, heart conditions may contribute to anxiety disorders, stress may be associated with sleep difficulties, and so on. This latter strategy may assist African Americans in feeling comfortable in self-reporting information specific to both physical and mental health. Indeed, there is some evidence to suggest African Americans are more likely to believe associated mental illnesses are influenced by physical impairments than Caucasians (Das et al. 2006).

Assessment of Ethnic Background To enhance service utilization we address culture early in the assessment process and discuss expectations of clients for therapy. We utilize the Semistructured Interview for Consideration of Ethnic Culture in Therapy Scale (SSICECTS), developed to assist providers in successfully determining the extent to which ethnic background should be addressed in treatment (Donohue et al. 2006). In a randomized controlled trial, Donohue et al. (2006) found that the SSICECTS could be utilized as a significant rapport builder in all ethnic groups examined. By acknowledging perceptions of ethnic cultural importance and empathizing with problems related to ethnic culture, those interviewed perceived their interviewers as having more knowledge and respect for their ethnic culture (as compared with control participants). In comparison with other ethnic groups, African Americans reported experiencing significantly more offensive remarks made toward them in their lifetime. Thus, there is a need to address the ethnic background of African Americans in therapy, and the SSICECTS is a psychometrically validated method of doing so.

Semistructured Interview for Consideration of Ethnic Culture in Therapy Scale (SSICECTS) commences with the interviewer assessing whether ethnic culture of clients is important to them (i.e., Ethnic Cultural Importance Scale, ECI). The question stems include: my culture is a big part of my everyday life, my ethnic culture is of great importance to me, there are things I like about my ethnic culture, and my ethnic culture should be addressed in therapy. Clients indicate the extent to which they agree with the item stem. For all endorsed ECI questions (i.e., the client agrees w/ stem content) the interviewer (1) validates that the client expressed agreement with the question (e.g., “You indicated that you agreed your ethnic culture is a big part of your life.”); (2) validates that content of the question is important (e.g., “I think it’s great that your ethnic culture is a big part of your life.”); (3) asks the client to elaborate on the respective response (e.g., “How is your ethnic culture a big part of your life?”); and (4) maintains pleasant and positive conversation (i.e., demonstrates agreement, elicits additional specific information). If disagreement is indicated by the client, the interviewer discloses that the client disagreed with the statement, and asks why disagreement was endorsed (e.g., “Help me to understand how your ethnic culture is not important to you?”). The interviewer then expresses empathy or understanding after the explanation is offered (e.g., “I see what you mean. Not having close family ties didn’t provide you opportunities to appreciate your ethnic cultural background.”), and assesses if the client was disappointed that disagreement was endorsed.

The next phase of the SSICECTS approach involves assessing problems experienced due to ethnicity (Ethnic Cultural Problems Scale, ECPS). The two question stems include, (1) experiencing offensive remarks due to ethnic culture, and (2) experiencing problems due to ethnic culture. If agreement is indicated, the interviewer then (1) discloses that the client agreed that others have made offensive comments about his or her ethnic culture; (2) reassures that others have also experienced similar comments; (3) asks the client to explain how the remarks were offensive; (4) empathizes with expressed concern; (5) asks how the offensive remarks affected the client; (6) either praises statements that suggest the individual has grown stronger because of the offensive remarks or provides empathy for statements that suggest the individual was negatively affected by the offensive remarks. If disagreement with the question stems is indicated, the interviewer determines if this was because others have not made remarks about the interviewee's ethnic culture or because the interviewee does not get offended.

The SSICECTS assists in understanding, directly from clients, how their ethnic background contributes to the therapeutic process, while establishing rapport. The interview only requires approximately 10–20 min to complete, making it a highly reliable and valid assessment tool that is likely to enhance rapport and assist African American substance abusing clients in feeling more comfortable reporting intimate issues.

Substance Use Assessment We highly recommend the utilization of valid and reliable self-report methods when assessing substance use among African Americans. The Time-Line Follow-Back interview (TLFB) consists of a month-by-month calendar that is utilized to determine substance use, and other problem behaviors, for up to 1 year prior to the day of assessment (Sobell et al. 1986). Important events (e.g., birthdays, graduations) are marked on the calendar to facilitate accurate recall prior to recording days of substance use. Other events may be additionally assessed, including school attendance, days employed, days incarcerated, days hospitalized, and frequency of HIV risk behaviors. Separate interviews with significant others may also be conducted to obtain frequency of client problem behaviors.

The client is queried during the TLFB interview to indicate antecedent stimuli that may influence substance use as reported in their calendars. Particular days in the week or time periods that suggest higher and lower rates of substance use are pointed out to provide a conceptual understanding of factors that appear to be maintaining substance use. The TLFB calendar may also be used to assess on-going substance use between treatment sessions. However, unless memory deficits are indicated, or there is an extended time period between treatment sessions, the calendar with memory anchor points is unnecessary. During treatment adult significant others should be taught to recognize signs of drug use (e.g., smell of marijuana, dilation of pupils, unsteady gait), thereby enhancing accuracy of their TLFB reports. Clients who abuse multiple substances may under-report "hard" drug use (e.g., cocaine) while disclosing marijuana or alcohol use. Therefore, it is very important to consider biological assessments of various substances, at least initially (Allen et al. 2009).

Substance use may be tested utilizing inexpensive on-site enzyme immunoassay tests purchased from reputable laboratories, such as Redwood Toxicology in California. However, screening may also be conducted by independent laboratories using, for example, SYVA Emit enzyme amino acid assay techniques with positive immunoassay screens verified using gas chromatography for alcohol and thin layer chromatography for all other substances. The panel of drugs varies, but often includes marijuana, cocaine, amphetamines, barbiturates, benzodiazepines, opiates, PCP, and methaqualone. Methyldioximethamphetamine ("crystal meth") and designer drugs are relatively less likely among African Americans, thus these panels may be unnecessary unless evidence of their use is suggested. Urinalysis is a reliable, quick, and cost-effective method for drug screening that correlates with TLFB reports retrospectively up to 6 months (Donohue et al. 2007). Alcohol use can be verified through a portable breathalyzer, up to 12 hours after ingestion (Allen et al. 2009).

Thus, whereas we recommend biological testing in the detection of drug abuse, and TLFB in the assessment of substance use frequency; there are other screening and comprehensive assessment measures for substance use and abuse that are widely used in the field and have been shown to be reliable and valid (please see Table 12.1). We will provide a brief overview of commonly used assessments of substance use and related problems in the following section. Lastly, while significant gains have been made in assessing substance use among the African American population, there still is a need for more empirical evidence supporting the use of the measures described below.

Overview of Selected Substance Use Screening Assessments

Addiction Severity Index (ASI) was designed to be administered as a semi-structured interview and is a multi-dimensional tool that not only as-

Table 12.1 At-a-glance summary table of substance use screening and assessment instruments

Assessment name	Disorder assessed	Instrument details	Recommended for use with African Americans
Addiction severity index (ASI) Teen addiction severity index (T-ASI)	Alcohol and other drugs	Semi-structured; 45–75 min clinical interview; adult and adolescent versions available	Yes
Alcohol use disorders identification test (AUDIT)	Alcohol	Brief screen for high risk alcohol use; administered by trained professionals; adults; 10 min	Yes
CAGE (CAGE)	Alcohol	Self-administered; clinician interview; less than 1 min	Mixed findings
Michigan alcoholism screening test (MAST)	Alcohol	Semi-structured; 5 min; adults and adolescents	Yes
Substance abuse subtle screening inventory (SASSI)	Alcohol and other drugs	Self-administered; 5–10 min; adults and adolescents	Mixed findings
Timeline followback (TLFB)	Alcohol and other drugs	Self-administered or by a clinician; 10–30 min; adults and adolescents	Yes

sesses for substance abuse, but also includes an overview of associated problems in seven areas (e.g., medical status, employment and support, legal status; McLellan et al. 2006). The ASI is a brief assessment tool which obtains objective and subjective data for each problem area within the last 30 days, as well as, over the client's lifetime. Objective data includes drug test results, and physical and mental health records. Subjective data includes client self-report severity ratings in seven domains including: medical status, employment and support, drug use, alcohol use, legal status, family and social status, and psychiatric status. It is the most widely used substance abuse assessment instrument in both clinical and research settings. The ASI may be utilized to obtain more detailed information about the nature and extent of substance use in African Americans because it has demonstrated high levels of reliability and validity across races and ethnicities (Brown et al. 1993).

CAGE (CAGE) is a screening tool that asks about lifetime alcohol or drug consumption (Ewing 1984). The CAGE is the most widely used alcohol detection measure and is commonly used in medical settings (Dhalla and Kopec 2007). The CAGE consists of four questions in a non-intimidating format and has a very quick administration time (Mayfield et al. 1974). The four questions of the CAGE form an acronym and consist of (1) Have you ever felt that you ought to *Cut* down on your drinking? (2) Have people *Annoyed* you by criticizing your drinking? (3) Have you ever felt bad or *Guilty* about your drinking? (4) Have you ever had a drink first thing in the morning to steady your nerves or to get rid of a hangover (*Eye-opener*)? (Mayfield et al. 1974). The CAGE has been shown to be a reliable and valid measure (Teitelbaum and Carey 2000). Some criticisms arise regarding the CAGE's performance in young adults (Dhalla and Kopec 2007) and other studies have found the validity of the CAGE to be better in men than women (Cherpitel 2000). There is limited research specifically addressing the use of the CAGE in African American populations, however in a review of alcohol screening questionnaires in women, the CAGE was found to perform better as a screening tool in African American women compared to Caucasian women (Bradley et al. 1998). Overall, the CAGE can be an effective tool in screening for alcohol use concerns. Its non-intimidating format and applicability to medical settings make the use of the CAGE a quick and easy tool for screening of African American substance use that is consistent with the recommendations to assess background information and elucidate the relationship between physical and mental health.

Alcohol Use Disorders Identification Test (AUDIT) is a screening tool used to identify hazardous and harmful alcohol use, abuse, and dependence (AUDIT; Saunders et al. 1993). The AUDIT was shown to be an effective alcohol-screening instrument for African Americans across the full spectrum of alcohol misuse including those that meet diagnostic criteria for alcohol abuse and dependence. The brief AUDIT-C is also an effective three-item screening test for detecting problematic alcohol use in African Americans. When compared to the CAGE, the AUDIT produced greater reliability and validity (Hays et al. 1995).

Michigan Alcoholism Screening Test (*MAST)* assesses alcohol abuse and is the most widely used assessment tool used for measuring lifetime alcohol abuse and related problems (MAST; Selzer 1971). The 25-item questionnaire can be administered rapidly and provides a valid and reliable assessment of lifetime alcohol use and alcohol-related problems (Gibbs 1983; Storgaard et al. 1994). Although there is a lack of empirical studies done solely with the African American population, one study's finding did reveal a significant association between African American college students who were adult children of alcoholics and the MAST (Rodney and Rodney 1996).

Substance Abuse Subtle Screening Inventory (SASSI) is a brief screening tool, easily administered and assesses alcohol and other drug use as well as related problems (SASSI; Miller 1985). In clinical settings, the SASSI results may help a clinician understand the client and allow them to provide effective feedback. The downside to the SASSI is that it is more time consuming to administer relative to the other instruments described and has also been shown to have lower predictive value in detecting African Americans substance use disorders (Peters et al. 2000).

Conclusion

As stated earlier, assessment is an on-going process in therapy. We believe that professionals continuously assess throughout treatment to ensure that the African American client's needs are appropriately addressed and effective outcomes are achieved. This chapter briefly discussed challenges African Americans face when substance use issues need addressing in a professional setting. We have highlighted our approach to overcome mistrust, stigma, and overall fear of seeking substance abuse treatment. We feel the assessment process is the initial step toward engaging African Americans who have been indicated to abuse substances, and establishing favorable first impressions. Offering a culturally sensitive, empathic, and supportive assessment experience, provides an opportunity to establish positive goals that hopefully will be achieved. We recognize that this chapter does not comprehensively address all important aspects of assessment in African Americans; however, we believe the reviewed approach is theoretically and empirically sound (and doable), permitting a first step in the substance abuse assessment of African Americans. We believe our approach is likely to enhance treatment retention rates among African Americans who at times may feel like their experiences are overlooked or not attempted to be understood. With sincere effort and focus on evidence-based assessment, such as the strategies offered in this chapter, we believe assessment of substance abuse in African Americans is likely to be successful.

Happy Assessing!

Appendix A

Psychiatric Diagnostic Assessment The structured clinical interview for DSM-IV (SCID-IV; First et al. 1992) may be used to assist in determining primary AXIS I mental health disorders listed in the DSM-IV-TR, including substance use disorder. This interview has acceptable psychometric properties (Azrin et al. 2001; Spitzer et al. 1992). Psychiatric diagnosis is required for insurance reimbursement and often needed to assure services (e.g., Medicaid). The SCID can be utilized to substantiate diagnoses of substance abuse and dependence, as well as comorbid diagnoses. It should be mentioned, however, that the SCID requires specialized training to administer, and requires about 2 hours to implement.

HIV Risk Assessment The HIV risk assessment battery (RAB) is utilized to assess HIV risk behaviors (Cohen et al. 1994). The RAB consists of eight items that assess HIV risk due to drug use practices, and nine items that assess risk of HIV due to sexual behavior. The information from this assessment can be used during intervention, as clients are informed that endorsed RAB items are associated with greater risk of HIV, and they are assisted in setting goals that are specific to managing these risk behaviors. As mentioned previously, African Americans are at high risk for HIV infection making the administration of the RAB very appropriate with this population.

Life Satisfaction Scale The life satisfaction scale (LSS) is a measure designed to assess client happiness in a 12 life domains, as well as, obtain an overall life satisfaction rating (Donohue et al. 2003). The 12 life domains assessed include friendships, family, school, employment/work, fun activities, appearance, sex life/dating, drug use, alcohol use, money/material possessions, transportation, and control over one's own life. Each domain is rated using a 0–100 % scale of happiness. The LSS maintains psychometric properties that are excellent and the brief nature of the instrument permits this tool to be implemented often throughout treatment as well (Donohue et al. 2003). For all domains that are rated less than 100 %, goals are determined by asking how the domain could be improved to 100 % satisfaction. Assessing life satisfaction in African Americans substance users is especially important, as a positive life orientation has been found to be a protective factor against future substance use in at-risk African-American youths (Kogan et al. 2005).

Family Environment Scale (FES; Moos and Moos 1994) The FES Conflict and Cohesion scales appear to be particularly useful measures in substance abuse (Santisteban et al. 2003). The Conflict scale measures the extent to which family members are perceived to argue and disagree, whereas the Cohesion scale measures the extent to which the family is perceived to be harmonious and "close." The role of family cohesion cannot be understated in African American households, as it has been shown to be significantly related with the age of onset for marijuana use for African American females (Elaine et al. 2008). In addition, a supportive family environment has been found to influence the alcohol use of adolescent African American males more so than their Caucasian counterparts (Watt and Rogers 2007). Psychometric properties are good.

Treatment Planning We utilize an innovative, standardized treatment planning technique that includes program orientation, goal-setting, and performance planning. First, the client is oriented to the specific aims and policies of the program. Our specific aims usually include avoidance of substance use, avoidance of HIV risk behaviors, mental health functioning, and positive relationships with others. Each measure is reviewed individually with an emphasis on client strengths. Focusing on strengths is essential, as studies show that clients in a strength-based, residential substance abuse treatment program find focusing on strengths to be the most useful part of treatment and beneficial to engagement (Harris et al. 2012). Treatment goals are then developed by reviewing goal-worthy items from each measure followed by continued assessment of the client's thoughts about the factors that influence each goal-worthy item to occur. Goals and solutions are then developed collaboratively by the client and clinician. The client is queried how each specific aim can be accomplished, and supportive feedback gathered from the assessment process is provided to assist in goal development. Whenever possible, positive, standardized solutions are provided to assist in goal identification. Lastly, performance programming is established, permitting clients and their significant others to choose intervention options from a list of evidence-based strategies. Each intervention offered by the program is described in detail and the perceived helpfulness of each intervention is solicited from both the client and supportive other(s). The interventions are then prioritized based on the perceived helpfulness to determine the order in which the interventions will be implemented. Of course, many of the reviewed assessment measures (usually a smaller battery of scales that were found to be problematic during the initial pre-treatment assessment) are administered periodically to assess progress in therapy.

References

Allen, D. N., Donohue, B., Sutton, G., Haderlie, M., & LaPota, H. (2009). Application of a standardized assessment methodology within the context of an evidence-based treatment for substance abuse and its associated problems. *Behavior Modification, 33,* 618–654.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.) Washington, DC: American Psychiatric Association.

Azrin, N. H., Donohue, B., Teichner, G., Crum, T., Howell, J., & DeCato, L. (2001). A controlled evaluation and description of Individual-cognitive problem solving and family-behavioral therapies in conduct-disordered and substance dependent youth. *Journal of Child and Adolescent Substance Abuse, 11,* 1–43.

Bolden, L., & Wicks, M. N. (2005). Length of stay, admission types, psychiatric diagnoses, and the implications of stigma in African Americans in the nationwide inpatient sample. *Issues in Mental Health Nursing, 26,* 1043–1059.

Bradley, K. A., Boyd-Wickizer, J., Powell, S. H., & Burman, M. L. (1998). Alcohol screening questionnaires in women. *JAMA: The Journal of the American Medical Association, 280,* 166–171.

Breslau, J., Kendler, K. S., Su, M., Gaxiola-Aguilar, S., & Kessler, R. C. (2005). Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. *Psychological Medicine, 35,* 317–327.

Breslau, J., Aguilar-Gaxiola, S., Kendler, K. S., Su, M., Williams, D., & Kessler, R. C. (2006). Specifying race–ethnic differences in risk for psychiatric disorder in a USA national sample. *Psychological Medicine, 36,* 57–68.

Brown, L. S., Alterman, A. I., Rutherford, M. J., Cacciola, J. S., & Zaballero, A. R. (1993). Addiction Severity Index scores of four racial/ethnic and gender groups of methadone maintenance patients. *Journal of Substance Abuse, 5,* 269–279.

Brunswick, A. F., Lewis, C. S., & Messeri, P. A. (1991). A life span perspective on drug use and affective distress in an African-American sample. *Journal of Community Psychology, 19,* 123–135.

Caetano, R. (2003). Alcohol-related health disparities and treatment-related epidemiological findings among Whites, Blacks, and Hispanics in the United States. *Alcoholism: Clinical and Experimental Research, 27,* 1337–1339.

Centers for Disease Control and Prevention. (2012). *New HIV infections in the United States*. http://www.cdc.gov/nchhstp/newsroom/docs/2012/HIV-Infections-2007-2010.pdf. Accessed 15 Jan 2014.

Cherpitel, C. J. (2000). A brief screening instrument for problem drinking in the emergency room: The RAPS4. *Journal of Studies on Alcohol and Drugs, 61,* 447–449.

Claus, R. E., & Kindleberger, L. R. (2002). Engaging substance abusers after centralized assessment: Predictors of treatment entry and dropout. *Journal of Psychoactive Drugs, 34,* 25–31.

Cohen, E., Navaline, H., & Metzger, D. (1994). High-risk behaviors for HIV: A comparison between crack-abusing and opioid-abusing African-American women. *Journal of Psychoactive Drugs, 26,* 233–241.

Dhalla, S., & Kopec, J. A. (2007). The CAGE questionnaire for alcohol misuse: A review of reliability and validity studies. *Clinical & Investigative Medicine, 30,* 33–41.

Das, A. K., Olfson, M., McCurtis, H. L., & Weissman, M. M. (2006). Depression in African Americans: Breaking barriers to detection and treatment. *The Journal of Family Practice, 55,* 30–39.

Donohue, B., Azrin, N. H., Lawson, H., Friedlander, J., Teichner, G., & Rindsberg, J. (1999). Improving initial session attendance of substance abusing and conduct disordered adolescents: A controlled study. *Journal of Child & Adolescent Substance Abuse, 8,* 1–13.

Donohue, B., Teichner, G., Azrin, N., Weintraub, N., Crum, T. A., Murphy, L., & Silver, N. (2003). Initial reliability and Validity of the Life Satisfaction Scale for Problem Youth in a Sample of Drug Abusing and Conduct Disordered Youth. *Journal Of Child And Family Studies, 12,* 453–464.

Donohue, B., Hill, H., Azrin, N. H., Cross, C., & Strada, M. (2007). Psychometric support for contemporaneous and retrospective youth and parent reports of adolescent marijuana use frequency in an adolescent outpatient treatment population. *Addictive Behaviors, 32,* 1787–1797.

Donohue, B., Strada, M. J., Rosales, R., Taylor-Caldwell, A., Hise, D., Ahman, S., Lefforge, N. L., Kopof, M., Devore, G., Soares, B., Radkovich, B., & Laino, R. (2006). The semistructured interview for consideration of ethnic culture in therapy scale: Initial psychometric and outcome support. *Behavior Modification, 30,* 867–891.

Elaine, D., Kerry, G., & Heather, R. (2008). Long-Term Patterns of Drug Use Among an Urban African-American Cohort: The Role of Gender and Family. *Journal of Urban Health, 85,* 250–267.

Ewing, J. A. (1984). Detecting alcoholism. *JAMA: the journal of the American Medical Association, 252,* 1905–1907.

Fendrich, M., Johnson, T. P., Wislar, J. S., Hubbell, A., & Spiehler, V. (2004). The utility of drug testing in epidemiological research: Results from a general populations survey. *Addiction (Abingdon, England), 99,* 197–208.

Fiscella, K., Franks, P., Doescher, M. P., & Saver, B. G. (2002). Disparities in health care by race, ethnicity, and language among the insured: Findings from a national sample. *Medical Care, 40,* 52–59.

Freimuth, V. A., Crouse-Quinn, S., Thomas, S. B., Cole, G., Zook, E., & Duncan, T. (2001). African Americans' views on research and the Tuskegee Syphilis study. *Social Science & Medicine, 52*(5), 797–808.

Gibbs, L. E. (1983). Validity and reliability of the Michigan Alcoholism Screening Test: A review. *Drug and Alcohol Dependence, 12,* 279–285.

Grant, B. F., & Harford, T. C. (1995). Comorbidity between DSM-IV alcohol use disorders and major depression: Results of a national survey. *Drug and Alcohol Dependence, 39,* 197–206.

Grosenick, J. K., & Hatmaker, C. M. (2000). Perceptions of the importance of physical setting in substance abuse treatment. *Journal of Substance Abuse Treatment, 18,* 29–39.

Harris, N., Brazeau, J. N., Clarkson, A., Brownlee, K., & Rawana, E. P. (2012). Adolescents' experiences of a strengths-based treatment program for substance abuse. *Journal Of Psychoactive Drugs, 44*(5), 390–397.

Hasin, D. S., Stinson, F. S., Ogburn, E., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: Results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry, 64,* 830–842.

Hays, D., Merz, F., & Nicholas, R. (1995). Response burden, reliability, and validity of the CAGE, Short MAST, and AUDIT alcohol screening measures. *Behavior Research Methods, Instruments, & Computers, 27,* 277–280.

Horton, J., Compton, W., & Cottler, L. B. (2000). Reliability of substance use disorder diagnoses among African-Americans and Caucasians. *Drug and Alcohol Dependence, 57*(3), 203–209.

Jacobson, J. O., Robinson, P. L., & Bluthenthal, R. N. (2007). Racial disparities in completion rates from publicly funded alcohol treatment: Economic resources explain more than demographics and addiction severity. *Health Services Research, 42*(2), 773–794.

Johnson, T., & Fendrich, M. (2005). Modeling Sources of self-report bias in a survey of drug use epidemiology. *AEP, 15,* 381–389.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62,* 617–627.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19.

Kim, M. T., & Hill, M. N. (2003). Validity of self-report of illicit drug use in young hypertensive urban African American males. *Addictive Behaviors, 28,* 795–802.

Kogan, S. M., Luo, Z., Murry, V., & Brody, G. H. (2005). Risk and Protective Factors for Substance Use Among African American High School Dropouts. *Psychology Of Addictive Behaviors, 19,* 382–391.

Kranke, D., Guada, J., Kranke, B., & Floersch, J. (2012). What do African American youth with a mental illness think about help-seeking and psychiatric medication?: Origins of stigmatizing attitudes. *Social Work in Mental Health, 10,* 53–71.

Lloyd-Jones, D., Adams, R. J., Brown, T. M., Carnethon, M., Dai, S., De Simone, G., & Wylie-Rosett, J. (2010). Heart disease and stroke statistics 2010 update: A report from the American Heart Association. *Circulation, 121,* 46–215.

Mayfield, D., McLeod, G., & Hall, P. (1974). The CAGE questionnaire: Validation of a new alcoholism screening instrument. *The American Journal of Psychiatry, 131,* 1121–1123.

McLellan, A. T., Cacciola, J. C., Alterman, A. I., Rikoon, S. H., & Carise, C. (2006). The Addiction Severity Index at 25: Origins, contributions and transitions. *The American Journal on Addictions, 15,* 113–124.

Menke, A., Orchard, T. J., Imperatore, G., Bullard, K. M., Mayer-Davis, E., & Cowie, C. C. (2013). The Prevalence of Type 1 Diabetes in the United States. *Epidemiology, 24,* 773–774.

Mensch, B. S., & Kandel, D. B. (1988). Underreporting of substance use in a national longitudinal youth cohort: Individual and interviewer effects. *Public Opinion Quarterly, 52,* 100–124.

Miller, G. A. (1985). *The substance abuse subtle screening inventory (SASSI) manual*. Springville: SASSI Institute.

Milligan, C. O., Nich, C., & Carroll, K. M. (2004). Ethnic differences in substance abuse treatment retention, compliance, and outcome from two clinical trials. *Psychiatric Services, 55,* 167–173.

Minsky, S., Vega, W., Miskimen, T., Gara, M., & Escobar, J. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. *Archives of General Psychiatry, 60,* 637.

Moos, R. H., & Moos, B. S. (1994). *Family environment scale manual*. Palo Alto: Consulting Psychologists Press.

Obasi, E. M., & Leong, F. T. L. (2009). Psychological distress, acculturation, and mental health-seeking attitudes among people of African descent in the United States: A preliminary investigation. *Journal of Counseling Psychology, 56,* 227–238.

Perera-Diltz, D. M., & Perry, J. C. (2011). Screening for adolescent substance-related disorders using the SASSI-A2: Implications for nonreporting youth. *Journal of Addictions & Offender Counseling, 31,* 66–79.

Peters, R. H., Greenbaum, P. E., Steinberg, M. L., Carter, C. R., Ortiz, M. M., Fry, B. C., & Valle, S. K. (2000). Effectiveness of screening instruments in detecting substance use disorders among prisoners. *Journal of Substance Abuse Treatment, 18,* 349–358.

Richardson, J., Fendrich, M., & Johnson, T. P. (2003). Neighborhood effects on drug reporting. *Addiction (Abingdon, England), 98,* 1705–1711.

Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E., Burke, J. D. Jr., & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of general psychiatry, 41,* 949.

Rodney, H. E., & Rodney, L. (1996). An exploratory study of African American collegiate adult children of alcoholics. *Journal of American College Health, 44,* 267–272.

Rosenthal, D. A., & Berven, N. L. (1999). Effects of client race on clinical judgment.

Rehabilitation Counseling Bulletin, 42, 243–264.

Santisteban, D., Coatsworth, J., Perez-Vidal, A., Kurtines, W., Schwartz, S., LaPerriere, A., & Szapocznik, J. (2003). Efficacy of brief strategic family therapy in modifying Hispanic adolescent behavior problems and substance use. *Journal of Family Psychology, 17,* 121–133.

Saunders, J. B., Aasland, O. G., Babor, T. F., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption- II. *Addiction (Abingdon, England), 88,* 791–804.

Selzer, M. L. (1971). The Michigan alcoholism screening test: The quest for a new diagnostic instrument. *American Journal of Psychiatry, 127,* 1653–1658.

Shavers, V. L., Lynch, C. F., & Burmeister, L. F. (2002). Racial differences in factors that influence the willingness to participate in medical research studies. *Annals of Epidemiology, 12,* 248–256.

Sobell, M. B., Sobell, L. C., Klajner, F., Pavan, D., & Basian, E. (1986). The reliability of the timeline method of assessing normal drinker college students' recent drinking history: Utility for alcohol research. *Addictive Behaviors, 11,* 149–162.

Spitzer, R. L., Williams, J. B., Gibbon, M., & First, M. B. (1992). The structured clinical interview for the DSM-III-R (SCID): I. History, rationale, and description. *Archives of General Psychiatry, 49,* 624–629.

Storgaard, H., Nielsen, S. D., & Gluud, C. (1994). The validity of the Michigan alcoholism screening test (MAST). *Alcohol and Alcoholism, 29,* 493–503.

Strada, M. J., Donohue, B., & Lefforge, N. L. (2006). Examination of ethnicity in controlled treatment outcome studies involving adolescent substance abusers: A comprehensive literature review. *Psychology of Addictive Behaviors, 20,* 11–27.

Substance Abuse and Mental Health Services Administration (2009). Treatment Episode Data Set (TEDS): 1997-2007. National Admissions to Substance Abuse Treatment Services, (DASIS Series: S-47, DHHS Publication No. (SMA) 09-4379). Rockville, MD.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different: Theory and practice* (2nd ed.). Oxford: Wiley.

Taxman, F. S., & Bouffard, J. A. (2003). Substance abuse counselors' treatment philosophy and the content of treatment services provided to offenders in drug court programs. *Journal of Substance Abuse Treatment, 25,* 75–84.

Teitelbaum, L. M., & Carey, K. B. (2000). Temporal stability of alcohol screening measures in a psychiatric setting. *Psychology of Addictive Behaviors, 14,* 401.

Thompson, V. L. S., Bazile, A., & Akbar, M. (2004). African Americans' perceptions of psychotherapy and psychotherapists. *Professional Psychology: Research and Practice, 35,* 19–26.

Treolar, C., & Holt, M. (2006). Deficit models and divergent philosophies: Service providers' perspectives on barriers and incentives to drug treatment. *Drugs: Education, Prevention and Policy, 13,* 367–382.

Watt, T. T., & Rogers, J. M. (2007). Factors contributing to differences in substance use among Black and White adolescents. *Youth Society, 39,* 54–71.

Williams, D. R., & Jackson, P. B. (2005). Social sources of racial disparities in health. *Health Affairs, 24,* 325–334.

Zhang, A. Y., & Snowden, L. R. (1999). Ethnic characteristics of mental disorders in five U.S. communities. *Cultural Diversity & Ethnic Minority Psychology, 5,* 2134–2146.

Assessing Eating Pathology in African Americans

13

Shanette M. Harris

African American Eating Pathology

In recent years increased research attention has been focused on the topic of clinical eating disorders (EDs) and dysfunctional eating among African Americans. Despite a growing number of publications, the research continues to centralize the results for European Americans, particularly European American females. Consequently, an unclear picture of empirical findings related to the eating attitudes, symptoms, and diagnoses of African Americans has emerged. These findings show that a contradictory and mixed body of studies dominate the literature and obscure preventive, diagnostic, and interventional implications for clients, therapists, and counselors. Several researchers report lower rates of clinical EDs and dysfunctional eating behaviors among African-heritage samples than European American samples (Crago et al. 1996; Halpern et al. 1999; Kraeplin 2011; Napolitano and Hime 2011; O'Neill 2003; Wildes et al. 2001). However, some studies that have examined eating behaviors for adult community females (e.g., Striegel-Moore et al. 2000; Wilfley et al. 1996), college undergraduate females (e.g., Franko et al. 2007), as well as adolescent females (e.g., Johnson et al. 2002; Kelly et al. 2011) provide evidence of similar rates of binge eating and bulimia between these two groups. More recently, researchers report that compulsive or binge eating and binge eating disorder (BED) have a high prevalence rate among African-heritage women (Taylor et al. 2007, 2013). African American women also appear to suffer from night eating syndrome (NES) associated with overweightedness and obesity (Jarosz et al. 2007). Although the prevalence of anorexia is lower than bulimia, subclinical levels of bulimia, any binge eating (ABE), and BED there is evidence that African-heritage American females experience anorexia nervosa (AN) and subthreshold symptoms of this disorder (Fernandes et al. 2010; Hoek and van Hoeken 2003; Marques 2011; Taylor et al. 2007; Taylor et al. 2013). Some research even shows that African-heritage women have a higher rate of EDs and eating symptomatology than European Americans. Marques et al. (2011) compared prevalence of EDs across ethnic minority groups and found that both 12-month and lifetime prevalences of BN were greater in Latinas and African Americans. The rate of ABE was also higher in Latinas, Asian Americans, and African Americans, while 12-month prevalence was higher in Latinos (men and women) and African Americans (men and women). In addition, all three minority groups reported more binge eating symptoms than non-Latina whites.

Ironically, race/ethnic differences in risk factors predictive of reported EDs have impeded the advancement of explanatory models that can

---

S. M. Harris (✉)
Department of Psychology, University of Rhode Island,
420 Chafee Social Sciences Center,
Kingston, RI 02881, USA
e-mail: sharris@mail.uri.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_13,

accommodate the diversity of variables that produce and maintain eating disturbances among African-descended ethnic groups. Researchers have relied on factors found to predict disordered eating behavior among European American females to understand the presence and absence of ED symptoms and behaviors among African-heritage women. Although there are obvious commonalities given the acculturation level of the majority of African-heritage Americans, the population is best understood as heterogeneous and composed of multiple ethnic groups. A sociocultural paradigm which holds that socialization and enculturation contribute to disordered eating behavior is applicable to African-heritage individuals but has not consistently incorporated African heritage worldviews.

Cultural values and norms have been theorized to underlie differences in the eating behaviors and symptoms of race/ethnic groups but culture-based risk factors have not been thoroughly examined. However, research studies with African-heritage females have been conducted on body image attitudes, culture-specific physical attractiveness standards, racial attitudes, ethnic identification, acculturative stress, and exposure to popular culture images and icons representative of beauty and health that operate through the media. These variables interact with gender role norms, family qualities, personality attributes, and oppressive experiences as a function of race and gender to produce multiple and complex body perspectives and relationships with food. Several studies demonstrate evidence of high appearance esteem and body satisfaction, body area satisfaction, high fitness and appearance evaluation, low drive for thinness but preference for heaviness, and low rates of weight and shape modifications among African-heritage or black females (e.g., Cash et al. 2004; Baugh et al. 2010, Chao et al. 2008; Crago et al. 1996; Grabe and Hyde et al. 2006; Harris 1994, 1995, 2001, 2006; Roberts et al. 2006; Sabik et al. 2010). These race/ethnic dissimilarities make the European American emphasis on the attainment of a thin body and consequent body dissatisfaction less relevant to a comprehensive model of eating pathology for African-heritage Americans, specifically African American females.

Assessment and African Americans

EDs have been constructed as psychological and mental health concerns that impact white females of Western culture, primarily in the USA. Observations of early theorists and researchers (e.g., Boskind-White 1987; Bruch 1966) that White females of upper socioeconomic status sought treatment and were hospitalized for EDs, particularly AN, resulted in failure to consider females of other racial/ethnic backgrounds and males. African Americans were marginalized from initial empirical research despite case studies that showed evidence of EDs (e.g., Robinson and Anderson 1985; Silber 1986). The invisibility of African Americans in research studies has contributed to their notable absence in the conceptualization and development of evaluation and diagnostic assessment methods. The development, standardization, and validation of interviews and instruments to assess eating pathology are steeped in the empirically established risk factors found to predict disordered eating for European Americans. Ideally, a similar percentage of African-heritage individuals should be included in standardization or norm samples during the development of interviews or instruments before employing them in clinical settings with African-heritage clients. However, the majority of the instruments and interviews developed to assess EDs have used predominantly or *only* European-heritage individuals. Typically any reference to race/ethnicity and cultural factors are made as an afterthought. For example, a review of the literature reveals that the majority of published studies with African Americans do not provide internal consistency data for self-report instruments. The use of measures primarily conceptualized with research related to eating dysfunctions of European American females and standardized with the same groups have the potential to yield biased and erroneous interpretative results when applied to individuals of other cultural worldviews.

Confirmation of initial evidence that African-heritage Americans are at risk for clinical EDs and disturbed eating behaviors and attitudes makes the standardization and validation of assessment methods for this population imperative. This chapter presents interviews and self-report instruments that have been used frequently in the ED research. These two types of assessment focus on screening (i.e., identification of those with high levels of eating pathology for further assessment), diagnosis, treatment planning, and evaluation of treatment outcomes. Assessment topics that require consideration relate directly to the symptoms of the disorders reflected in the DSM-IV-TR and symptom pictures presented in empirical studies. These topics include: body weight, binge eating and compensatory behaviors, concern with body weight and shape, dietary restraint, body image disturbance, and psychoemotional discomfort (Anderson et al. 2004). Each dimension is core to the ED experiences of European-heritage females and integral to the assessment methods described in the chapter. They represent the foundation upon which the interviews and self-report instruments have been conceptualized. Given that clear and consistent differences between European- and African-heritage females have been found for all dimensions but binge eating and compensatory behaviors, it is essential that researchers empirically examine the psychometric properties and validity of the interviews and instruments described. Limited information about the psychometric adequacy of instruments that are used in studies and clinical practice with African heritage women compounds existing discrepancies in the disordered eating literature.

Structured and Semistructured Interviews

Interviews are the primary method to gather data from a client in order to make a clinical diagnosis. Interviews provide an occasion to obtain more detailed information and to observe the client, which makes them more reliable for diagnostic classification decisions than for some other forms of assessment. Information can be obtained by using unstructured (common method of assessment with unknown psychometric properties (see Crowther and Sherwood 1997 and Peterson 2005), and semistructured and structured interviews. Several clinical interviews have been developed and are often used to assess ED pathology. Five are discussed in this chapter: The Eating Disorder Examination (EDE), children's version of the EDE (ChEDE), Clinical Eating Disorder Rating Instrument (CEDRI), Interview for the Diagnosis of Eating Disorders (IDED)-IV, and the Structured Interview for Anorexic and Bulimic Disorders for Expert Ratings (SIAB-EX).

The Eating Disorder Examination (EDE)

The EDE (Cooper and Fairburn 1987; Fairburn and Cooper 1993; Fairburn et al. 2008) is described as "the gold standard of ED assessment" (Berg et al. 2012, p. 428) against which other ED assessments are validated. The EDE is a semistructured interview that is designed to assess the attitudinal/cognitive and behavioral features of eating dysfunctions and generates a global score and four subscale scores. The interview is comprehensive and detailed which makes it lengthy to complete and relies on the clinical decision of the investigator which requires prior training sessions. The four EDE subscales that assess cognitive symptoms of EDs include: dietary restraint (i.e., reflects attempts to restrict food intake to influence weight or shape), eating concern (i.e., reflects the degree of concern about eating), and the shape concern and weight concern subscales (i.e., reflect concern about shape and weight) and the influence of these concerns on self-evaluation. Each item is rated on a seven-point forced-choice scale format (0–6) such that higher scores indicate higher frequency or greater severity. The EDE interview is used for assessment and diagnosis. Use of the EDE for assessment purposes requires that items refer to 28 days preceding the interview date. When used for diagnostic purposes, the items are modified to evaluate information for a 3-month time period consistent with DSM-IV diagnostic criteria (except items for BED which were modified to include questions to assess diagnostic requirements for a 6-month period). Behaviors are rated on the number of days on which they occurred and on the number of episodes. These include objective bulimic episodes (OBEs, i.e., binge eating defined as unusually large amounts of food with a subjective feeling of loss of control) and number of days on which these occurred for the previous month, subjective bulimic episodes (SBEs, i.e., subjective view of loss of control while eating food that is not viewed as large), and objective overeating episodes (OOEs, i.e., overeating without perceived loss of control). The schedule is composed of several sections that examine demographics, weight, menstrual history, dieting, binge eating, weight control, exercise, history of abuse, psychiatric, chemical, medical, and medication history, along with a medical checklist, and an assessment of family and social history. The EDE can be used with clients who present with any of the four EDs included in the DSM-IV-TR (4th edition, text revision; APA 1994) including AN, BN, eating disorder not otherwise specified (EDNOS), and the binge eating provisional diagnosis. Data obtained from the EDE can also be analyzed in two different ways. First, subscale scores and behavior frequency scores can be used as dimensional variables. In addition, Fairburn and Cooper (1993) presented algorithms that can be determined based on clients' diagnostic status to provide categorical information.

Evidence supports the reliability (interrater, test–retest, internal consistency, long term recall of symptoms) and validity (criterion-oriented, construct, discriminative that compares ED cases to no cases, and convergence with similar constructs and daily food records) of the EDE (Berg et al. 2012). The lowest reliability coefficients for internal consistency have been found for the four subscales in samples with BED or community-based samples, whereas the highest internal consistencies are reported for samples of females with full and subthreshold AN and BN (Berg et al. 2012). However, published studies have not examined the interrater reliability of scores on items that assess laxative and diuretic use or excessive exercise. Construct validity for the factor structure of the EDE also shows that of three studies conducted, none have successfully replicated the original four-factor model of the interview (Berg et al. 2012). In one study, a two-factor model was the best fit with the first factor similar to restraint and the second factor, a combination of the other three subscales. A second study reported a three-factor model (body dissatisfaction, shape/weight overevaluation, and dietary restraint) (Grilo et al. 2010). Other studies report a one-factor model (weight and shape concern) (Bryne et al. 2010; Wade et al. 2008). Both the two- and three-factor models suggest that body dissatisfaction (i.e., shape/weight) is associated with normative discontent but overvaluation of shape and weight is associated with more pathology (Grilo et al. 2010).

Low reliability coefficients for BED may have implications for African Americans who tend to experience binge eating and BED at higher rates than AN and BN. There is an obvious absence of data on the EDE with African American adults and adolescents for both males and females. Research is especially needed on the psychometric properties of the EDE in clinical and nonclinical samples of African heritage individuals. A version of the EDE has also been developed for children and adolescents (i.e., ChEDE) (Bryant-Waugh et al. 1996).

Child Eating Disorder Examination (ChEDE)

The Child Eating Disorder Examination 12.0 (ChEDE; Bryant-Waugh et al. 1996) is a semistructured, interviewer-based measure designed to assess the major attitudinal and behavioral correlates of EDs in children between 8 and 14 years of age. Similar to the adult version of the EDE interview, the ChEDE yields four subscale scores (i.e., restraint (R), eating concern (EC), weight concern (WC), shape concern (SC), and a global score that assesses the overall severity of ED psychopathology. In addition, the measure generates two forms of overeating: OBEs and SBEs. The OBE form involves an obvious large amount of food consumption, followed by a sense of loss of control over eating. SBEs involve consumption of a large amount of food that is seen as excessive by the individual but is not unambiguously large, and loss of control is present. Objective operating episodes (OOs) involves consumption of an unambiguously large amount of food in the absence of loss of control over eating: objective overeating (OO) and subjective overeating (SO). OO refers to episodes without loss of control but a large amount of food is consumed. SO refers to episodes without loss of control but the amount of food is not rated as large by the interviewer. The ChEDE also includes diagnostic items that can be used to obtain a clinical diagnosis of an ED. Unlike the EDE, a card sort task to address overvaluation of body shape and weight is included in the ChEDE. Respondents also complete a diary of events over the previous 4 weeks to help orient them to the period being examined. The wording was modified to increase comprehension for younger children.

Studies of the ChEDE for youth with EDs have shown global and subscale scores comparable to those observed in adult patients with EDs (Bryant-Waugh et al. 1996). Watkins et al. (2005) reported that the ChEDE had good basic psychometric properties. The measure was found to have sensitive discriminative validity related to group classification. Cronbach alpha coefficients were R (α=0.80), EC (α=0.91), WC (α=0.90), and SC (α=0.88). Similar to the EDE, as of this date, studies have not been published that examine the psychometric properties of the ChEDE with African-descent children.

Clinical Eating Disorder Rating Instrument (CEDRI)

The Clinical Eating Disorders Rating Instrument (CEDRI; Palmer et al. 1987) is a semistructured interview designed to be administered by a clinician to assess the behaviors and beliefs related to AN and BN with adults and adolescents. The interviewer rates eating attitudes, behaviors, and general psychopathology. The interview has been shown to discriminate between the clinically eating disordered and dieters without EDs (Palmer et al. 1996). Adequate interrater reliability has been demonstrated (Palmer et al. 1987). There is little published research on this interview with European American samples and an absence of research with African-heritage samples.

Interview for Diagnosis of Eating Disorders-IV (IDED)

The Interview for Diagnosis of Eating Disorders (IDED; Kutlesic et al. 1998) was originally developed by Williamson (1990) using DSM-II-R criteria and has undergone three revisions. The fourth version or current form is the IDED-IV (Kutlsic et al. 1998). This is a quick method to screen for EDs with adults and adolescents. The IDED is a semistructured interview designed to differentially diagnose AN, BN, EDNOS, and BED according to DSM-IV criteria. The interview consists of 20 items (symptoms) and uses a five-point Likert scale. Items are summed for each subscale and a total score calculated based on summing the subscale scores. If all ratings within a subscale are 3 or higher, a diagnosis of BN, AN, or BE is determined. Kutlesic et al. (1998) present four studies that discuss psychometric properties of the interview. It has been shown to have good internal consistency with Cronbach α=AN (0.75), BED (0.96), and BN (0.75). This more recent version inquires about demographics and ED history. There seems to be limited psychometric data available for the IDED for European American samples and the interview has not been empirically examined with black or African-heritage samples.

Structured Interview for Anorexic and Bulimic Disorders of Expert Ratings (SIAB-EX)

The Structured Interview of Anorexic and Bulimic Disorders of Expert Ratings (SIAB-EX; Fichter et al. 1991) is the most recent (third) version of the SIAB-EX and consists of 87 items. Twenty of these items are used to provide a diagnosis according to ICD-ID or DSM-IV criteria. The majority of the items are coded on a 5-point scale ranging from 0 (symptom/problems not present) to 4 (symptoms/problem very severely present). The content includes all topics relevant to EDs and symptoms of emotional disturbance associated with EDs (e.g., anxiety, depression, and substance abuse). The interview can be used with all individuals who are thought to have an ED. The interview can also be used in epidemiologic research (i.e., SIAB self-rating version). Computer algorithms are able to compute sum scores and diagnostic classifications. The SIAB-EX covers a wider range of symptoms than the EDE (Fichter and Quadflieg 2001) and follows the DSM-IV more strictly. Current expression of symptoms, most symptoms ever experienced in the past, and lifetime expression of symptoms are assessed. Each item has clearly defined symptoms and criteria and case examples used for scoring. Unlike the EDE, the factor analytic approach is used by this schedule from the start (Fichter and Quadflieg 2001). The SIAB-EX consists of six subscales: body image and thinness ideal, general psychopathology/past and general psychopathology and social integration (GenPsySoc; current), sexuality and social integration (past) and sexuality (Sex; current), bulimic symptoms, inappropriate compensatory behaviors to counteract weight gain, fasting, substance abuse, (compensatory behavior), and atypical binges. A manual that defines each item and item score can be used to train clinical interviewers. The SIAB-EX provides a computer algorithm for the diagnosis of BED according to DSM-IV criteria and other ED disturbances. Each item has clearly defined symptoms and criteria and case examples. However psychometric data for African heritage populations are nonexistent.

Self-Report Instruments

Several types of self-report measures have been published that can be used for varied purposes. Similar to interviews, self-report instruments can provide information that can assist clinicians with diagnostic decisions. They can also assess behavior or symptom changes as a function of an intervention/treatment or change in symptomatology across time. Frequently viewed as a weakness of self-report is the finding that more dysfunction is endorsed on these instruments than in face-to-face interview contexts. However, self-report instruments may benefit diagnosis and treatment planning by allowing the respondent to avoid anxiety, shame, and embarrassment that can interfere with the provision of forthright responses. Table 13.1 contains self-report instruments used to assess EDs that are discussed in this section.

The Eating Disorder Examination Questionnaire (EDE-Q)

The Eating Disorder Examination Questionnaire (EDE-Q; Fairburn and Beglin 1994) is a self-report questionnaire derived from the interview-EDE (described above; Fairburn and Cooper 1993). The current version of the EDE-Q is the EDE-Q 60 (Fairburn and Beglin 2008). The 33-item EDE-Q is basically the same as the EDE. This instrument also examines ED symptoms and attitudes for the past 28 days and contains the same four subscales: restraint (R), eating concern (EC), shape concern (SC), and weight concern (WC). The subscale and total scores are based on averages from 0 to 6, with higher scores reflecting greater pathology. Items related to the frequency of symptoms of binge eating and compensatory behaviors for the past 28 days are also included. However the EDE-Q does not require special training and is more efficient for assessment in clinical and nonclinical settings.

Research findings show good convergence between subscales of this instrument and the EDE interview ($r=0.68$–0.76) but participants score higher on the EDE-Q than the EDE interview (Berg et al. 2011). This agreement with the EDE is in measurement of attitudes and compensatory behavior for both eating-disordered and noneating-disordered samples. These findings suggest that the EDE-Q as compared to the EDE interview overestimates the severity of eating psychopathology (Black and Wilson 1996; Fairburn and Beglin 1994). Allen et al. (2011) evaluated the

Table 13.1 Self-report instruments for EDs

Assessment name	Disorder assessed	Recommendations and/or relevant research findings
EDE-Q	ED symptoms and attitudes	Best for diagnostic classification; low reliability for BED; eating concern (EC) subscale good screening instrument. High scores may not indicate disturbed behavior. Need psychometric data
EAT-40 and EAT-26	Eating attitudes and symptoms	Effective as a screening instrument in nonclinical settings; useful as an outcome measure in clinical populations. May not assess same construct for African and European heritage samples. Need for psychometric data
ChEAT	Eating attitudes and dieting	Designed for normal and underweight children; cutoff score may not be appropriate for all age levels
EDDS	AN, BN, and BED	Useful for clinical and nonclinical samples; brief and useful for diagnosing AN, BN, and BED; does not use the word "binge". May not assess same construct for race/ethnic groups
EDI, EDI-2, and EDI-3	Traits related to AN and BN	EDI-3 useful as a general or nonclinical screening instrument; interoceptive deficits scale is best predictor and DT no longer necessary. BD subscale needs further research for construct validity
BULIT-R	Bulimic symptomatology	Good reliability. Strong in the assessment of BED. Evidence or comparability for race/ethnic groups
BES	Features of binge eating	Identifies nonbinging but weak at identifying those engaged in binging; more useful to identify psychopathology related to binge eating rather than the disorder. Appears to assess different constructs for African and European samples. BULIT-R measures binge eating
TFEQ, TFEQ-21, and TFEQ-RI8V2	Restraint	Need more psychometric data with different samples. Atlas et al. (2002) found that restraint was less stable over time for African American females
MEBS	Disturbed eating (AN, BN, BED, and EDNOS)	Can use with adults and children as young as 10 years old. Lack of psychometric data and few studies with African heritage groups
SCOFF	Eating Disorders	A clinical cutoff has not been established. Research that examines psychometric properties is necessary. Moderate effectiveness as a screening instrument; detailed questions should follow positive results for the SCOFF. Need psychometric research with African heritage samples
QEWP, QEWP-R, QEWP-A, and QEWP=P	BED	Psychometric data is needed with African heritage samples

factor structure of the instrument in community/university and clinical samples and found that the original four-factor model was not supported. The best fit was a brief one-factor model consisting of eight weight and shape concern items for community and eating disordered samples. They concluded that the EDE and EDE-Q are best at diagnosis and do not empirically assess degree or severity of ED symptoms well. Vander Wal et al. (2011) reported that only the Eating Concern (EC) subscale of the EDE-Q was sufficient as a brief screening tool in the detection of

BED among populations with high rates of EDs. They suggest that the EC subscale can be useful in office settings, with telephone screenings, and in epidemiological studies. Hrabosky and Grilo (2007) reported good internal validity for Black and Hispanic samples: R (0.84 and 0.79), EC (0.80 and 0.70), SC (0.92 and 0.90), and WC (0.84 and 0.83). Stojek and Fischer (2013) reported a Cronbach alpha for EDE-Q-restraint of $\alpha = 0.80$ at T1 of their study and $\alpha = 0.88$ at T2 for an African American sample.

More recently, Kelly et al. (2012) provided EDE-Q norms for undergraduate Black women ($N = 395$), subscale scores, and clinical cutoff data which they compared to previously published norms for two white samples (i.e., a white undergraduate and a community sample). The Cronbach alpha coefficients showed high reliability and were comparable with those of previous publications (R) $\alpha = 0.82$, (EC) $\alpha = 0.79$, (WC) $\alpha = 0.88$, (SC) $\alpha = 0.91$, and (global pathology) $\alpha = 0.91$. Using a clinical cutoff score of ≥ 4 as indicative of clinically significant eating pathology the findings were similar to those of the white undergraduate female sample with the exception of WC which was significantly higher for the black sample. On the subscales, the black sample was similar to the white community sample as there were no differences between the groups. However, compared to the white undergraduates the black sample reported significantly less R and EC. There were no differences on the other subscales or global pathology. The black sample differed from both white samples on eating disordered behaviors. Compared to their white undergraduate counterparts, fewer reported engaging in regular subjective binge eating and compensatory behaviors (e.g., self-induced vomiting and laxative or diuretic misuse). A similar picture emerged between black undergraduates and the white community sample. As the black sample reported less participation in regular subjective or objective binge eating and compensatory behaviors (e.g., self-induced vomiting or excessive exercise). Interestingly, all participants endorsed EDE-Q scores at a similar percentage, however, black female undergraduates reported less dysfunctional eating. This finding suggests that even when black females endorse eating concerns at a high level such scores may not necessarily indicate problematic eating or compensatory behaviors. From a psychometric perspective, the test seems highly reliable although this cutoff score may not accurately predict the EDs and symptomatology of black female undergraduates. The assessment of culturally-relevant variables along with the EDE-Q might provide more insight into EDs for this group (e.g., acculturative stress and ethnic identity).

Eating Attitude Test-40 (EAT-40) and Eating Attitude Test-26 (EAT-26)

The original Eating Attitude Test-40 (EAT-40) was developed by Garner and Garfinkel (1979). The development of the EAT-40 instrument involved the administration of the initial item set to two samples of females with anorexia and two control groups of females to identify attitudes and behaviors associated with AN (Garner and Garfinkel, 1979). The instrument uses a 6-point Likert scale format with the following response options: (1) always, (2) very often, (3) often, (4) sometimes, (5) rarely, and (6) never. The EAT has been shown to have moderate to good reliability (coefficient $\alpha = 0.79$) for an anorexic sample (Garner and Garfinkel 1979) and a high degree of internal consistency ($r = 0.94$) with the identification of seven factors (Garner and Garfinkel 1979). Items 1, 18, 19, 23, and 39 are scored 6 = 3 points; 5 = 2 points; 4 = 1 point; and 3, 2, or 1 = 0 points. The remaining items are scored: 1 = 2 points; 2 = 2 points; 3 = 1 point, and 4, 5, or 6 = 0 points. The scores for each item differ from one another and the total score is the sum of each item. A total score of 30 was established as the cutoff for identifying individuals symptomatic with anorexia (Garner and Garfinkel 1979). That is, a total score higher than 30 is interpreted as indicative of an anorexic disorder. The EAT has been found to differentiate between nonanorexics and anorexics and to identify college females with subclinical eating problems (e.g., Button and Whitehouse 1981; Garner and Garfinkel 1980).

The Eating Attitude Test-26 (EAT-26; Garner et al. 1982) was developed based on two females samples: female patients seen for consultation related to AN ($N=160$) and a female comparison sample of freshman and sophomore college females of ($N=140$). The 40 items of the EAT were factor analyzed for the AN sample and yielded three factors: (a) dieting (relating to avoidance of high calorie food and preoccupation with thinness); (b) bulimia and food preoccupation (thoughts about food and bulimia); and (c) oral control (related to self-control of eating and pressure from others to gain weight). The 14 items not loading on the 3 factors were eliminated resulting in the EAT-26. Answers are also based on a six-point Likert scale (i.e., always = 3, usually = 2, often = 1, sometimes = 0, rarely = 0, never = 0). The question "enjoying trying new rich foods" is the only reverse-scored item.

A cut-off score of 20 on the EAT correctly classified a similar proportion the AN group and comparison group (i.e., 84.9% based on total score) and is the typical threshold cutoff for the presence of significant ED symptomatology. Intercorrelations showed that the EAT-26 is highly predictive of the total EAT-40 (i.e., 0.98). Factor I of the EAT-26 has the highest correlation with the total scale ($r=0.93$) followed by factor II ($r=0.64$) and factor III ($r=0.60$). The internal consistency reliability for the EAT-26 was high (i.e., $\alpha=0.90$). The authors proposed that the instrument can be used as an outcome measure in clinical populations or as a screening instrument in non-clinical settings. However the EAT-26 is often used as a screening instrument and has the same three subscales as the EAT-40.

Both the EAT-40 and EAT-26 have been widely used as screening instruments for disturbed eating and have been translated into several languages (e.g., Spanish, Portuguese, and Greek) and validated with different cultural groups (e.g., Brazilian, Greek, Malaysian, Turkish, Portuguese, and Spanish). The instruments have also been used with African American females, however, a review of the literature does not reveal a psychometric evaluation of either EAT measure with African-heritage Americans. Like the EDE and ChEDE, a child version of the EAT has also been developed for use with children.

Children's Eating Attitude Test (ChEAT)

The Children's Eating Attitude Test (ChEAT; Maloney et al. 1989) is a 26-item measure of children's eating attitudes and dieting behaviors was primarily designed for normal weight children and underweight children with AN and BN. The scale uses a six-point Likert scale. The ChEAT was designed based on the Eating Attitudes Test (EAT; Garner et al. 1982), which was developed and tested on adolescents and adults. Mahoney et al. (1988) modified the items of the EAT to make them more easily understood by children as young as 8 years old. Example questions include "I eat diet foods" and "I am terrified of being overweight." Validity testing of the ChEAT supports recoding scores such that the least three symptomatic responses (never, rarely, and sometimes) are recoded as 0, with often = 1, usually = 2, and the most symptomatic score always coded = 3. The total ChEAT score can range from 0 to 78. The cutoff point is 20. High scores (i.e., 20 points or more) indicate the presence of ED symptomatology. The ChEAT has shown adequate to good reliability. Among adolescent girls, internal consistency coefficients range from 0.87 to 0.89 (Maloney et al. 1989). A 3-week test-retest reliability coefficient of $r=0.81$ has been demonstrated (Maloney et al. 1988). Concurrent validity of the ChEAT has been examined in several studies. ChEAT scores have significant, positive correlations with weight management ($r=0.36$; $p<0.001$) and body dissatisfaction ($r=0.39$; $p<0.001$) (Smolak and Levine 1994). ChEAT scores also have significant negative correlations with body areas satisfaction (Cash 1997; Sancho et al. 2005). Smolak and Levine's (1994) examination of the psychometric properties of the ChEAT with 308 middle school girls had adequate reliability but a 23-item version yielded higher reliability. Factor analysis of the 26-item version revealed factors similar to the EAT but yielded an additional fourth factor (i.e., restricting and purging). Vander Wal and Thomas (2004) reported an α of 0.74 for African American girls and 0.82 for Hispanic girls with a fourth- and fifth-grade inner city public school sample.

There is some evidence to suggest that the cutoff score of 20 may need further examination with culturally different populations. Escoto et al. (2010) called for more research related to a cutoff score of 20 because of differences noted between the percentage of Mexican children identified and the percentages detected in samples of other researchers. Sancho et al. (2005) used 17 or higher as an "experimental" cutoff in the Spanish version of the ChEAT and requested a lower cutoff in order to detect more students with eating problems (i.e., low sensitivity). The original cutoff for the EAT has been shown not to be accurate for all age groups and for certain ages a cutoff score of 25 is more appropriate (Erickson and Gerstle 2007). There is an absence of published psychometric information with African heritage children for the ChEAT.

Eating Disorder Diagnostic Scale (EDDS)

The Eating Disorder Diagnostic Scale (EDDS; Stice et al. 2000) is a recently developed brief self-report measure for diagnosing AN, BN, and BED according to DSM-IV diagnostic (American Psychiatric Association 1994) criteria. The scale comprises 22 items that include Likert scores, dichotomous scores, frequency scores, and open-ended questions. The first four items, attitudinal symptoms of AN and BN in the past 3 months (e.g., fear of fatness) are measured on a seven-point scale (0–6). The next four items assess frequency of uncontrollable consumption of a large amount of food with attention on the days per week for a 6-month period (BED) and number of times per week for the past 3 months (BN). The third set of items measure the frequency of behaviors used to compensate for binge eating over the past 3 months (e.g., vomiting, laxative or diuretic use, fasting, and excessive exercise). The final items relate to weight and height, birth control pills and the menstrual cycle. A diagnosis is made of the three types of EDs and an ED symptom score that differentiates eating disordered from healthy controls is generated.

The EDDS is useful to assess eating problems for clinical and nonclinical samples. Interestingly, the instrument does not use the word "binge" which may be especially helpful for African-heritage respondents who interpret the word in a different way than the behavior (Fernandez et al. 2006). The measure can be used for diagnostic purposes and for ongoing assessment of symptomatology. The initial validation yielded test-retest reliability for the composite of 0.87 and Cronbach's $\alpha = 0.89$. Mitchell and Mazzeo (2004) reported a strong internal consistency score (Cronbach $\alpha = 0.89$). An examination of the psychometric features of the EDDS revealed high internal consistency for the full sample (full sample, $\alpha = 0.94$), clinical sample ($\alpha = 0.86$), and nonclinical sample ($\alpha = 0.87$). Internal consistencies of the symptom composite scores two weeks from baseline for clinical and nonclinical groups were ($\alpha = 0.80$) and ($\alpha = .86$), respectively. Excellent test-retest reliability based on scores made at baseline and 2 weeks later showed a kappa score of 0.86 for the eating disordered versus noneating disordered group. The accuracy rate was 0.95 showing almost a perfect relationship from baseline to the second measurement. Criterion validity was based on the degree of agreement between EDDS diagnoses and interview diagnoses with the EDE. Kappa agreement between the EDE and EDDS was 0.89. Sensitivity for AN and BN was above 0.95 but the sensitivity for BED was low (0.57). The positive predictive value was also lower for BED (0.80) than AN (1.00) and BN (0.86). Adequate convergent validity with general (e.g., depression and body attitudes) and eating pathology (R, EC, WC, and SC) was shown for the EDDS composite score (Krabbenborg et al. 2012). Research with nonclinical samples has shown evidence of the validity and reliability of the EDDS (Stice et al. 2000; Stice et al. 2004)—good test–retest reliability, content validity, criterion validity, and predictive validity. A symptom composite cutoff score of 16.5 was used based on the sample. Research on this instrument focuses only on European and American non-Latino white female clients and nonclinical samples. Studies should increase the generalizability of the use of the scale by examining the validity and reliability of the EDDS for African American men and women.

Eating Disorder Inventory (EDI-1, EDI-2, and EDI-3)

The Eating Disorders Inventory (EDI; Garner et al. 1983) was designed to assess behavioral and psychological traits related to AN and BN. The first version of the EDI consisted of 64 items based on published research and empirical data and grouped into eight subscales: drive for thinness (DT), BN, interpersonal distrust (ID), interoceptive awareness (IA), perfectionism (P), maturity fears (MF), body dissatisfaction (BD), and ineffectiveness (IE). Three of the subscales assess attitudes toward eating, weight, and body shape (DT, BN, BD) and the other five assess psychological characteristics related to EDs (ID, IA, ID, MF, IE). All three versions of the EDI require the respondent to rate each item on a six-point Likert scale using “always,” “usually,” “often,” “sometimes,” “rarely,” or “never.” These response options are coded into four response categories (0–3). A score of zero is assigned to the response farthest in the direction of “symptomatic.” The subscale scores can be added to obtain a total score or each subscale can be used separately. The total score is 192 and the cutoff score is 42 or less for the eight subscales. Some researchers propose that the EDI is useful as a screening instrument in the general or nonclinical population. However others suggest that the EDI yields low specificity (i.e., does not differentiate EDs from other disorders), which can hinder the screening potential. Yet, the EDI successfully differentiates individuals with EDs from those free from EDs.

The second version or EDI-2 was published in 1991 and items increased from 64 to 91. These items include the original eight subscales and the addition of 27 new items grouped into three subscales that measure general features of EDs including: asceticism (AS), impulse regulation (IR), and social insecurity (SI) (Garner 1991). Psychometric weaknesses of the EDI include: an imbalance in positively and negatively coded items, differences in the number of items assigned to subscales which can lead to acquiesence bias for culturally different respondents, and an absence of factor analytic support or only moderate support for 8 factors on the EDI-1 and 11 factors on the EDI-2 (Podar and Allik, 2009).

The most recent revision of items from the EDI-2 is the EDI-3 (Garner 2004). The EDI-3 consists of the same 91 items on the EDI-2, including the same three subscales of eating disordered symptoms (DT, BN, BD). The reliability scores from eating disordered samples are strong (Cronbach α=.90-.91; test-retest, r=.98) (Garner, 2004; Wildes et al., 2010). However, criticisms of the factor structure of the EDI-2 resulted in a new factor analysis of the same scores providing new subscales consistent with current research and theory on EDs. The EDI-3 consists of general psychological trait subscales: low self-esteem (LSE), personal alienation (PA), interpersonal insecurity (II), interpersonal alienation, interoceptive deficits, emotional dysregulation (ED), perfectionism (P), asceticism (AS), and maturity fear (MF). In this recent version, three response style indicators have been added (Garner 2004). The EDI-3 uses the six-choice format of the EDI-2 but the scores were changed from a 0–3 to 0–4 format with the objective of enhancing the psychometric properties of the instrument for nonclinical samples. The discriminative validity and internal consistency are good and better than in the original development (Clausen et al. 2011). A confirmatory factor analysis basically supported a grouping of eating problems in two general factor scores: a risk component and a component to assess pathological disturbance. Clausen et al. (2011) based their model fit on 90 item scores, rather than the 12 subscale sum scores which yielded a better fit than Garner (2004). Only one new item is in the EDI-3 compared to the EDI-2 and the bulimia subscale is a strong predictor of a BN diagnosis. According to the receiver operating characteristic (ROC) analyses of Clausen et al. (2011), interoceptive deficits is the best predictor across all ED diagnostic groups, followed by low self-esteem and personal alienation. They report that current use of drive for thinness (DT) is no longer necessary. Several studies have used subscales from versions of the EDI (e.g., DT, BD) with African American females. However more psychometric research is needed.

The Bulimia Test—Revised (BULIT-R)

The Bulimia Test—Revised (BULIT-R; Thelen et al. 1991) is a 36-item, multiple-choice, self-report, Likert-type scale used to measure the degree of bulimic symptomatology according to the DSM-III-R (American Psychiatric Association, 1987). Twenty-eight of the items are used to calculate a total score and there are no subscale scores on the instrument. All items are presented in a five-point, forced-choice format. Items are scored from 1 to 5 and a total score is obtained by summing across the 28 items. The scores can range from 28 (no symptoms) to 140 (elevated symptoms), and 104 is considered a cutoff for classification of BN. Those who score above this level are categorized as more likely to receive a diagnosis of BN in an interview than those who score below this level. Thelen et al. (1991) reported that scores on the BULIT-R distinguish between those with and without BN. They also found that the 2-month estimated temporal stability of scores on the scale was 0.95. Vander Wal et al. (2011) reported a specificity (100 %), sensitivity (96 %), positive predictive value (94 %), and negative predictive value (100 %) for the BULIT-R in the assessment of BED for a sample of overweight or obese individuals with BED and a control group with 26 non-BED individuals.

In a comparison of Caucasian and African American samples, Atlas et al. (2002) used structural equation modeling and found that a model specifying a one-factor structure to the BULIT fit the data and the instrument was comparable for both groups of females (comparative fit index, CFI = 0.98, nonnormed fit index, NNFI = 0.97; indexes range from 0 to 1 and values above 0.90 are indicative of a good fit).

Fernandes et al. (2006) examined the factor structure of the BULIT-R with African American, Asian American, Latino American, and Caucasian American college females ($N = 2671$). The Cronbach alpha ($\alpha = 0.92$) for the African American subsample ($N = 192$) shows strong reliability for the BULIT-R, and there were no differences on the clinical cut off of 104 for the groups. However, African Americans scored significantly lower than the Caucasian subsample on the BULIT-R, demonstrating that African Americans reported fewer bulimic attitudes and behaviors than did Caucasians. The five factors presented in Thelen's (1991) research were tested with confirmatory factor analysis (CFA) and found to be a poor fit for the data. Exploratory factor analyses were used to examine the factor structure of the instrument for each ethnic group. A six-factor solution was best for the four groups. Within each group, six factors explained almost 60 % of the variance in scores for each group (59 % for African Americans). A unique speed/amount of eating factor emerged for African and Asian Americans. Phrases or items that did not include the words "binge" or "binge eat" loaded on this factor, but questions that loaded on the binge factor did use these words. The discovery of two factors: one that uses the words binge or binge eating and one that does not but has a similar meaning is important for the language used and how questions are interpreted. The control items loaded on the body image/control factor. The four purging behaviors loaded on two or three factors for African Americans: vomiting/laxatives and diuretics. Unlike the other three groups, an exercise factor was not revealed. In general the between group differences need further examination but the reliability score, size of the population and overall substantiation of binge behavior, body image, extreme weight loss, and compensatory behaviors were measured for African Americans.

Binge Eating Scale (BES)

The Binge Eating Scale (BES; Gormally et al. 1982) is a 16-item self-report measure of binge eating that was developed to identify behavioral and cognitive features of binge eating in obese individuals. Each item is associated with three or four statements with different weights and a final score that ranges from 0 to 46. The scale can be used to identify binge eaters, evaluate binging severity, and assess the impact of treatment. The scale can also be used to screen for BED (Greeno et al. 1995). The BES has been found to discriminate effectively among individuals with no (scores of 17 or lower), moderate (scores between 18 and 26), and severe binge eating problems (scores of 27 or higher; Gormally et al. 1982). Developers of this measure report that it yields internally consistent scores (Cronbach's $\alpha = 0.85$) (Gormally et al. 1982).

Timmerman (1999) examined the validity of the BES with overweight binge eaters to assess severity of binge eating using food records. Severity of a binge was measured by total number of binge calories, binge episodes, and binge days over 28 days. BES scores were moderately associated with subjective and objective binge eating severity but were not related to total calorie intake. The author concluded that the BES was unable to differentiate between subjective and objective binges; an important distinction for diagnostic and treatment purposes. Greeno et al. (1995) also reported that the BES may not accurately identify binge eaters when binging or nonbinging classification is determined by the EDE semistructured interview. In their research, the BES correctly identified 92.9 % of nonbinging participants but only 51.8 % of those identified as binging with the EDE. Despite this low agreement, the authors suggest that the BES may be a useful screening tool for those with severe binge eating behaviors. Other findings of low agreement between the BES and ED instruments (e.g., Gladis et al. 1998) suggest that the BES may identify psychopathology related to binge eating rather than evidence of the specific disorder even with high scores. However, the scale is inadequate for diagnostic decisions which always requires an interview.

A Portuguese version of the BES has been found to be effective or useful in the assessment of binge ED in a clinical sample of obese Brazilian women. Using a cutoff score of 17 the scale was compared to the Structured Clinical Interview for the DSM-IV. A sensitivity of 97.8 %, specificity of 47.7 %, positive predictive value of 66.7 %, and negative predictive value of 95.3 % were found. Test–retest reliability after 121 days was 0.66 measured by kappa statistics and Cronbach alpha was 0.89 (Freitas et al. 2006). Webb and Hardin (2012) reported an internal consistency of $\alpha = 0.88$ for a sample of 134 first-time, first-year undergraduates of which 40 % were black/African American. Mitchell and Mazzeo (2004) reported a Cronbachs alpha of 0.91 for a nonclinical college sample with white and African American females.

Three-Factor Eating Questionnaire (TFEQ, TFEQ-21,TFEQ-R18V2)

The Three-Factor Eating Questionnaire (TFEQ; Stunkard and Messick 1985) previously known as the Eating Inventory (EI) was created to improve upon weaknesses of the Restraint Scale (RS). The TFEQ consists of 51 items that measure cognitive restraint (21 items), disinhibition (16 items), and hunger (14 items). Cognitive restraint seems to assess a dieting to disinhibition-of-restraint sequence (Marcus and Wing, 1983), and hunger relates to dieting but predicts binge eating (Stunkard and Messick 1985). Responses are scored 0–1 and summed. Higher scores indicate higher levels of restrained eating, disinhibited eating, and predisposition to hunger. Mazzeo et al. (2003) investigated the internal structure of the TFEQ or EI with confirmatory factor analysis and found that the original three-factor model that was proposed fit poorly with the data. Factor loadings on the TFEQ-R (i.e., cognitive restraint) were the most difficult to interpret. The TFEQ has been found to have good reliability, with coefficient alphas ranging from 0.79 to 0.84 for a sample of dieters and from 0.84 to 0.92 for nondieters (Stunkard and Messick 1985). The TFEQ has also been found to perform similarly for African American and Caucasian undergraduates (Atlas et al. 2002). Bardone-Cone and Boyd (2007) also found high reliability (>0.80) for the instrument at baseline and good reliability

(>0.76) at the second measurement 5 months later with the exception of the instability shown on subjective binge eating and dietary restraint for the African American sample. There was also evidence of convergent and discriminant validity for the TFEQ. For instance, dietary restraint measures were more highly related to each other than with bulimic symptom measures for both groups and measures of bulimic symptoms were more strongly related with each other than with dietary restraint.

The TFEQ-R21 is a more psychometrically sound version of the restraint subscale of the TFEQ that also comprises three subscales that assess cognitive restraint (tendency to control fool in order to modify shape and weight), uncontrolled eating (tendency to lose control over eating or when exposed to food stimuli), and emotional eating (propensity to engage in overeating when experiencing negative affect or mood). The instrument consists of 20 items rated to a four-point scale (i.e., definitely true/mostly true/mostly false/definitely false) and one one eight-point item. The scores are summed to provide scale scores and the raw scores are converted with a 0–100 scale. Higher scores indicate higher levels of cognitive restraint (CR), uncontrolled eating (UE), and emotional eating (EE).

The TFEQ-R18V2 is revised from the original 21-item instrument and comprises three subscales that measure cognitive restraint, uncontrolled eating, and emotional eating. Scores are derived from ratings of how true 18 statements are for respondents. The TFEQ-R18V2 has shown a stable or strong factor structure and good reliability. Three items were removed from the CR item set (CFI=0.91). A Cronbach alpha of 0.78–0.94 was shown (Cappelleri et al. 2009). The relation between CR and body mass index (BMI) seems to depend upon attributes of the sample (e.g., obese vs. nonobese; healthy vs. diabetes). Only one study by Atlas et al. 2002 has examined the reliability, validity, and factor structure of the TFEQ with an African American sample. Many more studies are needed with different ethnic groups within the African heritage population to gauge the relevance of these findings with revisions of this measure.

Minnesota Eating Behavior Survey (MEBS)

The Minnesota Eating Behavior Survey (MEBS; Klump et al. 2000; von Ransom et al. 2005) originally known as the Minnesota Eating Disorder Inventory, is a 30-item self-report instrument that uses four subscales to assess disturbed eating: body dissatisfaction, weight preoccupation, compensatory behavior, and binge eating. The scale was originally designed to provide indicators of behavioral and attitudinal symptoms associated with AN, BN, and BED associated with EDNOS diagnostic categories. Items assessing eating disordered cognitions and behaviors from the BD, BN, and DT subscales of the EDI (Garner et al. 1983) and two behavioral items from the IA subscale along with five additional items to assess compensatory behavior not already specific to the EDI constituted the item set. The language of the items were modified for comprehension by children as young as 10 years old. The directions ask respondents to circle “True” or “False” with “T” and “F” corresponding to each item. If the item is endorsed in the disordered direction the item receives one point and no point if answered in the nondisordered direction. Four items are reverse scored and the total MEBS score is derived from the sum of scores from the 30 items. Subscale scores are also derived from summing scores associated with each subscale. The response format was changed to a four-point scale (“definitely true,” “probably true,” “probably false,” or “definitely false”) 3 years after the original format. The MEBS has been shown to have internal consistency scores that range from 0.71 to 0.85 across the four subscales and moderate to high concurrent validity (von Ranson et al. 2005). The two primary cognitive/affective subscales are body dissatisfaction and weight preoccupation. The two primarily behavioral subscales are binge eating and compensatory behaviors (Klump et al. 2000; Marmorstein et al. 2007). A clinical cutoff score has not been determined for the MEBS. However, Harrell et al. (2009) examined family history of alcoholism disordered eating and depressive symptoms as predictors of college female alcohol concerns and found the sample showed a range of eating symptoms (i.e., range = 0–29; *M* = 8.43; *SD* = 6.43. Six percent indicated significant levels of disordered eating with MEBS total scores of 20 points or higher. In addition, 27 women or 9 % scored a 5 or higher on the binge eating subscale suggesting significant dysfunction for some in the sample. The MEBS showed high internal consistency in this college female sample ($\alpha = 0.89$). Reporting psychometric properties of the MEBS for a Canadian university female sample, von Ranson et al. (2007) reported a coefficient alpha of 0.92 for the total scale and coefficient alphas of 0.80 or better for three of the subscales but a low internal consistency for the compensatory behavior subscale ($\alpha < 0.60$). There is an absence of research on the psychometric properties of the MEBS. Validity, reliability, and factor structure information are needed for samples of adults and children of diverse race/ethnic groups.

SCOFF Interview/Questionnaire

The SCOFF questionnaire (SCOFF; Morgan et al. 1999) is a brief self-report instrument for ED screening that was developed in the UK. Five questions addressing the main symptoms of AN and BN were determined with focus groups comprised of adult female ED patients between 18 and 45 years of age who were referrals from a specialty clinic (*N* = 116; i.e., AN and BN diagnoses according to the DSM-IV criteria) and ED experts. College students free from EDs between 18 and 39 years old served as the control group (*N* = 96). Patients and members of the college control group were presented with the questions orally. They also completed the EDI and Bulimic Investigatory Test (BITE, a self-rating scale for BN). The acronym SCOFF is based on the letters in each of the five questions:

- Do you make yourself **S**ick because you feel uncomfortably full?
- Do you worry that you have lost **C**ontrol over how much you eat?
- Have you recently lost more than **O**ne stone (14 lb) in a 3-month period?
- Do you believe yourself to be **F**at when others say you are too thin?
- Would you say that **F**ood dominates your life?

A threshold or cutoff set at two or more questions for all five questions yielded a 100 % sensitivity (i.e., proportion of persons detected positive for the disorder) for AN (94–100 %) and BN (92.6–100 %), individually and when joined (AN and BN; 96.9–100 %) with a specificity (i.e., proportion of persons without diagnosis detected as negative) of 87.5 % for the control participants.

In a primary care assessment study, using the two or more question cutoff, the SCOFF detected all four cases of AN and BN and seven out of nine cases of EDNOS (Luck et al. 2002). A sensitivity of 84.6 % was found. Out of the 328 women who reported the absence of an ED the SCOFF indicated 34 false positives. The measure showed a specificity of 89.6 % (86.3–92.9 %), positive predictive value of 24.4 % (12.9–39.5 %), and negative predictive value of 99.3 % (97.6–99.9 %). The results suggest that the instrument is effective in assessment in primary care settings. The SCOFF may not be as suitable to detect patients who do not meet full criteria for AN or BN. The low positive predictive value stems from the low prevalence of EDs in the sample. However, Mond et al. (2008) compared the SCOFF and EDE-Q in a primary care sample of young women and found that the EDE-Q performed somewhat better in screening for ED cases (EDE-Q, sensitivity = 0.80, specificity = 0.80, and positive predictive value = 0.44) than the SCOFF (sensitivity = 0.72, specificity = 0.73, and positive predictive value = 0.35). The EDE-Q was also less vulnerable to the effects of body weight and age variation than the SCOFF. The authors recommend that any positive results with the SCOFF should follow with additional questions rather than immediate referral. Research also shows that the SCOFF is reliable in both oral interview and written formats, although participants showed higher scores with the written than the oral delivery, despite the order in which they were provided (Perry et al. 2002). The SCOFF has not been used with African heritage groups.

Questionnaire of Eating and Weight Patterns (QEWP, QEWP-R, QEWP-A, QEWP-P)

The Questionnaire of Eating and Weight Patterns (QEWP; Spitzer et al. 1992) was developed to assess aspects of BED as presented as a provisional diagnosis in the DSM-IV. The instrument consists of 12 stem items and several are followed with items to obtain greater detail. Most items are rated on a two-point scale with a "yes" and "no" response. Other items are rated on a five-point scale. Items that request "how often behaviors occur" are responded to with five options: less than 1 day/week, 1 day/week, 2 or 3 days/week, 4 or 5 days/week, and almost every day. Items that request information about "feeling bad" also have five response options: not bad at all, just a little bad, pretty bad, very bad, and very, very bad. The items are grouped into three diagnostic categories: no diagnosis, nonclinical binging, and BED. In a sample of 52 self-referred binge eaters and 52 control participants (19–75 years old), test–retest reliability after 3 weeks with 39 self-referred and 40 control group individuals showed a kappa of (0.58). The probability of binge eating for the BED group at 1 week and 3 weeks yielded a correlation of ($r=0.70$). At 1 week the sensitivity was 0.71 %, specificity was 0.69 %, and predictive efficiency value was 0.71 %. At 3 weeks, the sensitivity was 0.82 %, specificity was 0.63 %, and predictive efficiency value (0.74 %) (Nangle et al. 1994).

The QEWP-A and QEWP-P are modified versions of the QEWP.

The Adolescent Version–Questionnaire of Eating and Weight Patterns (QEWP-A)

The Adolescent Version of the Questionnaire of Eating and Weight Patterns (QEWP-A; Johnson et al. 1999; Johnson et al. 2000). This measure was adapted from the QEWP to assess binge eating in adolescents and consists of 13 items that relate to behavioral symptoms of BED. The modification involved combining two original items into one which reduced the item pool from 13 to 1. The language was also modified to make suitable for an adolescent population. Some response options are on a two-point scale, "yes"/"no". Other options are on a five-point scale. Adolescents are classified the same as with the QEWP (no diagnosis, nonclinical binge eating, and BED.). The test–retest reliability of the QEWP-A was determined in a study with 12–18 year-old males and females (Johnson et al. 2001). The reliability was more stable for males than females, 33 % of females changed from an initial classification of nonclinical binge eating to no diagnosis (Johnson et al. 2000). Concurrent validity was assessed by using the QEWP-A to predict depression and eating attitudes with the ChEAT. Participants classified with BED according to the QEWP-A show higher scores on a depression measure and eating attitudes measured with the ChEAT.

The Parent Version–Questionnaire of Eating and Weight Patterns (QEWP-P)

The Parent Version of the Questionnaire of Eating and Weight Patterns (QEWP-P; Johnson et al. 1999) is based on the QEWP and uses the same questions as the QEWP-A. The items were modified for parents to refer to their children (e.g., During the past 6 months did your child…). Their seems to be a low concordance between the QEWP-A and QEWP-P (Johnson et al. 1999). The two measures showed good agreement for an absence of disturbed eating behavior but were not found to be concordant as regards type of binge eating, overeating episodes, or compensatory weight control behaviors in a 6 month period. Steinberg et al. (2004) also found poor agreement for the QEWP-P and QEWP-A in a study with 142 overweight and 121 normal weight children between 6 and 12 years old. Children categorized by parents as engaging in binge eating, having greater body adiposity, eating cognitions, body dissatisfaction, and parent reported problems than children not engaging in overeating or simple overeating according to the QEWP-P but children categorized by their own reports (QEWP-A) as engaging in no overeating, simple overeating or binge eating did not differ in body composition or in eating and general psychological problems. Jones and Cook-Cotton (2013) reported a Cronbach alpha for the QEWP-A of 0.70 for African American adolescents.

Recommendations and Conclusion

This chapter has presented a sample of the numerous interview and self-report instruments that have been developed to assess eating pathology with adults, adolescents, and children. The instruments vary in reliability, validity, and whether empirical data yield a factor structure consistent with the original theoretical and empirical framework. Clinical interviews are most appropriate for diagnostic purposes but are costly, time consuming, and require carefully trained interviewers, and therefore, may be less suitable for screening needs, evaluation outcomes, or clinical research studies. Likewise, screening instruments are brief, economical, and can be used in settings that reach a broad number of individuals who may need assistance with EDs (e.g., primary care settings) but should not be used for diagnostic classification without an interview.

The overwhelming majority of interviews and instruments have not been empirically evaluated in terms of reliability, validity, and factor structure relevant for African heritage populations. A review of the literature reveals that between 2002 and 2012, approximately seven studies have been published that examine the psychometric properties of self-report instruments and interviews that are used with samples of African heritage. Of this limited number of studies, it appears that the self-report instruments vary in their reliability and accuracy with groups other than those for which they were normed. It is hoped that this chapter will encourage researchers to move to conduct research on ED assessment with African-descended Americans. Given proposed diagnostic changes in the new DSM-V, it is important that any forthcoming modifications in ED assessments associated with the new diagnoses incorporate psychometric information for this population. At this point, however, researchers can benefit by including additional scales that have been shown to assess cultural factors associated with eating disordered behavior among African-descended Americans (e.g., acculturative stress, general stress, stage of racial identity, ethnic identity, and race and gender microaggressions). Researchers should also incorporate other variables such as socioeconomic status and objective indicators of BMI in evaluations of the factor structure, reliability, and validity of instruments. It is also important that samples of men, children, and noncollege populations are included in studies that examine the psychometrics characteristics of current instruments. For instance, Kelly et al. (2012) recently found that several of the instruments that researchers use most often to assess eating attitudes and behaviors with African heritage samples do not assess the same underlying constructs. CFA did not support invariance of factor loadings for the EDDS and EDI (BD and DT subscales), EAT-26, and BES suggesting that these instruments measure different constructs for African-heritage Americans and European-heritage Americans. African American women also scored differently than their European heritage peers on the ED measures that CFA supported invariance of the loadings (e.g., BULIT-R). Interestingly, the majority of the 18 studies in O'Neill's (2003) analysis used BD and DT subscales of the EDI and the EAT-26 to assess disturbed eating behavior and attitudes. As we move to increase our awareness of ED risk beyond members of mainstream American culture, it is important to begin with an examination of the constructs that we measure to determine their relevance for other cultural groups.

References

Allen, K. L., Byrne, S. M., Lampard, A., Watson, H., & Fursland, A. (2011). Confirmatory factor analysis of the Eating Disorder Examination-Questionnaire (EDE-Q). *Eating Behaviors, 12,* 143–151. doi:10.1016/j.eatbeh.2011.01.005.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders*. (DSM-III-R) (3rd ed., revision), Washington, DC: APA.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders. (DSM-IV)* (4th ed.), Washington, DC: APA.

Anderson, D. A., Lundgren, J. D., Shapiro, J. R., & Paulosky, C. A. (2004). Assessment of eating disorders: Review and recommendations for clinical use. *Behavior Modification, 28*(6), 763–782.

Atlas, J. G., Smith, G. T., Hohlstein, L. A., McCarthy, D. M., & Kroll, L. S. (2002). Similarities and differences between Caucasian and African American college women on eating and dieting expectancies, bulimic symptoms, dietary restraint, and disinhibition. *International Journal of Eating Disorders, 32,* 326–334. doi:10.1002/eat.10081.

Bardone-Cone, A. M. & Boyd, C. A. (2007). Psychometric properties of eating disorder instruments in Black and White young women: Internal consistency, temporal stability, and validity. *Psychological Assessment, 19*(3), 356–362.

Berg, K. C., Peterson, C. B., Frazier, P., & Crow, S. J. (2012). Psychometric evaluation of the Eating Disorder Examination and Eating Disorder Examination-Questionnaire: A systematic review of the literature. *International Journal of Eating Disorders, 45*(3), 428–438.

Black, C. M., & Wilson, G. T. (1996). Assessment of eating disorders: Interview versus questionnaire. *International Journal of Eating Disorders, 20,* 43–50.

Boskind-White, M., & White, Jr., W. C. (1987). *Bulimarexia: The binge/purge cycle*. New York: Norton.

Bruch, H. (1966). Anorexia Nervosa and it's differential diagnosis. *Journal of Nervous and Mental Disease, 141,* 555–566.

Bryant-Waugh, R., Cooper, P., Taylor, C., & Lask, B. (1996). The use of the eating disorder examination with children: A pilot study. *International Journal of Eating Disorders, 19,* 391–398.

Button, E. J., & Whitehouse, A. (1981). Subclinical anorexia nervosa. *Psychological Medicine, 11,* 509–516.

Bryne, S. M., Allen, K. L., Lampard, A. M., Dove, E. R., & Fursland, A. (2010). The factor structure of the Eating Disorder Examination in clinical and community samples. *International Journal of Eating Disorders, 43*(3), 260–265. doi:10.1002/eat.20681.

Cash, T. F. (1997). *The body image workbook: An 8-step program for learning to like your looks*. Oakland: New Harbinger Publications.

Cash, T. F., Morrow, J. A., Hrabosky, J. I. & Perry, A. A. (2004). How has body image changed? A cross-sectional investigation of college women and men from 1983 to 2001. *Journal of Consulting and Clinical Psychology, 72*(6), 1081–1089.

Cappelleri, J. C., Bushmakin, A. G., Gerber, R. A., Leidy, L. K., Sexton, C. C., Lowe, M.R., & Karlsson, J. (2009). Psychometric analysis of the Three-Factor Eating Questionnaire- R21: Results from a large diverse sample of obese and non-obese participants. *International Journal of Obesity, 33,* 611–620. doi:10.1038/ijo.2009.74.

Chao, Y. M., Pisetsky, B. A., Dierker, L. C., Dohm, F., Rosselli, F., May, A. M., & Striegel-Moore, R. H. (2008). Ethnic differences in weight control practices among U. S. adolescents from 1995 to 2005. *International Journal of Eating Disorders, 41,* 124–133.

Clausen, L., Rosenvinge, J. H., Friborg, O., & Rokkedal, K. (2011). Validating the Eating Disorder Inventory-3 (EDI-3): A comparison between 561 female eating disorders patients and 878 females from the general population. *Journal of Psychopathology and Behavioral Assessment, 33,* 101–110. DOI: 1007/s10862-010-9207-4.

Cooper, Z., & Fairburn, C. G. (1987). The Eating Disorders Examination: A semistructured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders, 6,* 1–8.

Crowther, J. H., & Sherwood, N. E. (1997). Assessment. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (2nd ed., pp. 34–49). New York: Guilford.

Erickson, S. J., & Gerstle, M. (2007). Developmental considerations in measuring children's disordered eating attitudes and behaviors. *Eating Behaviors, 8,* 224–235.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16*(4), 363–370.

Fairburn, C. G., & Beglin, S. J. (2008). Eating disorder examination questionnaire (EDE-Q 6.0). In Fairburn, C. G. (Ed.), *Cognitive behavior therapy and eating disorders*. New York: Guilford.

Fairburn, C. G., & Cooper, Z. (1993). The eating disorder examination. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (12th ed., pp. 317–360). New York: Guilford.

Fernandes, N. H., Crow, S. J., Thuras, P., & Peterson, C. B. (2010). Characteristics of Black treatment seekers for eating disorders. *International Journal of Eating Disorders, 43*(3), 282–285.

Fernandes, S., Malacrne, V. L., Wilfley, D. E., & McQuaid, J. (2006). Factor structure of the Bulimia Test-Revised in college women from four ethnic groups. *Cultural Diversity and Ethnic Minority Psychology, 12*(3), 403–419.

Fichter, M. M., Elton, M., Engel, K., Meyer, A, E., Mall, H., & Poustka, F. (1991). Structured interview for anorexia and bulimia nervosa (SIAB): Development of a new instrument for the assessment of eating disorders. *International Journal of Eating Disorders, 10*(5), 571–592.

Fichter, M. M., & Quadflieg, N. (2001). The structured interview for anorexic and bulimic disorders for DSM-IV and ICD-10 (SIAB-EX): Reliability and validity. *European Psychiatry, 16,* 38–48.

Franko, D.L., Becker, A.E., Thomas, J.T., & Herzog, D.B. (2007). Cross-ethnic differences in eating disorder symptoms and related distress. *International Journal of Eating Disorders, 40,* 156–164.

Freitas, S. R., Lopes, C.S., Appolinario, J. C., & Coutinho, W. (2006). The assessment of binge eating disorder in obese woman: A comparison of the binge eating scale with the structured clinical interview for the DSM-IV. *Eating Behaviors*, *7*(3), 282–289.

Garner, D. M. (1991). *Eating disorder inventory-2: Professional manual*. Odessa: Psychological Assessment Resources.

Garner, D. M. (2004). *EDI-3: Eating disorder inventory-3: Professional manual*. Odessa: Psychological Assessment Resources.

Garner, D. M., & Garfinkel, P. E. (1979). The Eating Attitudes Test: An index of the symptoms of anorexia nervosa. *Psychological Medicine, 9,* 273–279.

Garner, D. M., & Garfinkel, P. E. (1980). Sociocultural factors in the development of anorexia nervosa. *Psychological Medicine, 10,* 647–656.

Garner, D. M., Olmstead, M. P., Bohr, Y., Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric features and clinical correlates. *Psychological Medicine, 12,* 871–878.

Garner, D.M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders, 2*(2), 19–34.

Gladis, M. M., Wadden, T. A., Foster, G.D., Vogt, R.A., & Wingate, B. L. (1998). A comparison of two approaches to the assessment of binge eating in obesity. *International Journal of Eating Disorders, 23*(1), 17–26.

Gormally, J., Black, S., Daston, S., & Rardin, D. (1982). The assessment of binge eating severity among obese persons. *Addictive Behaviors, 7,* 47–55.

Greeno, C.G., Marcus, M. D., & Wing, R. R. (1995). Diagnosis of binge eating disorder: Discrepancies between a questionnaire and clinical interview, *International Journal of Eating Disorders, 17*(2), 153–160.

Grilo, C. M., Crosby, R. D., Peterson, C. B., Masheb, R. M., White, M. A., Crow, S. J., Wonderlich, S. A., & Mitchell, J. (2010). Factor structure of the Eating Disorder Examination interview in patients with binge-eating disorder. *Obesity, 18*(5), 977–981. doi:10.1038/ob.2009.321.

Halpern, C. T., Udry, J. R., Campbell, B., & Suchindran, C. M. (1999). Effects of body fat on weight concerns, dating and sexual activity: A longitudinal analysis of Black and White adolescent girls. *Developmental Psychology, 35*(3), 721–736.

Harrell, Z. A. T., Slane, J. D., & Klump, K. L. (2009). Predictors of alcohol problems in college women: The role of depressive symptoms, disordered eating, and family history of alcoholism. *Addictive Behaviors, 34,* 252–257. doi:10.1016/j.addbeh.2008.10.019.

Harris, S. M. (1994). Racial differences in predictors of college women's body image attitudes. *Women and Health, 21*(4), 89–104.

Harris, S. M. (1995). Family, self, and sociocultural contributions to body-image attitudes of African-American women. *Psychology of Women Quarterly, 19*(1), 129–145.

Harris, S. M. (2001). Body image and weight as correlates of eating behaviors for African American college females. *Journal of Social and Behavioral Sciences, 38*(1), 31–41.

Harris, S. M. (2006). Body image attitudes, physical attributes and disturbed eating among African American college women. *Race, Gender & Class, 13*(1–2), 46–58.

Hoek, H. W. & Hoeken, D. V. (2003). Review of the prevalence and incidence of eating disorders. *International Journal of Eating Disorders, 34*(4), 383–396.

Hrabosky, J. I. & Grilo, C. M. (2007) Body image and eating disordered behavior in a community sample of Black and Hispanic women. *Eating Behaviors, 8,* 106–114. doi:10.1016/j.eatbeh.2006.02.005.

Jarosz, P. A., Dobal, M. T., Wilson, F. L., & Schram, C. A. (2007). Disordered eating and food cravings among urban obese African American women. *Eating Behaviors, 8,* 374–381.

Johnson, W. G., Grieve, F. G., Adams, C. D., Sandy, J. (1999). Measuring binge eating in adolescents: Adolescent and parent versions of the Questionnaire of Eating and Weight Patterns. *International Journal of Eating Disorder, 26,* 301–314.

Johnson, W. G., Kirk, A. A., & Reed, A. E. (2000). Adolescent version of the Questionnaire of Eating and Weight Patterns: Reliability and gender differences. *International Journal of Eating Disorders, 29,* 94–96.

Johnson, W. G., & Rohand, K. J., & Kirk, A. A. (2002). Prevalence and correlates of binge eating in white and African American adolescents. *Eating Behaviors, 3*(2), 179–189.

Jones, L. A., & Cook-Cotton, C. (2013). Media and cultural influences in African-American girls' eating disorder risk. *Preventive Medicine*. dx.doi.org/10.5402/2013/319701.

Kelly, N. R., Cotter, W. W., & Mazzeo, S. E. (2011). Eating Disorder Examination Questionnaire (EDE-Q): Norms for Black women. *Eating Behaviors, 13,* 429–432. doi.org/10.1016/j.eatbeh.2012.09.001.

Kelly, N. R., Mitchell, K. S., Gow, R. W., Trace, S. E., Lydecker, J. A., Bair, C.E., Mazzeo, S. (2012). An evaluation of the reliability and construct validity of eating disorder measures in White and Black women. *Psychological Assessment, 24*(3), 608–617. doi:10.1037/a0026457.

Klump, K. L., McGue, M., & Iacono, W. G. (2000). Age differences in genetic and environmental influences on eating attitudes and behaviors in preadolescent and adolescent female twins. *Journal of Abnormal Psychology, 109*(2), 239–251. doi:10.1037//0021-843X.109.2.

Krabbenborg, M. A. M., Danner, U. N., Larsen, J. K., van der Veer, N., van Elburg, A. A., de Ridder, D. T. D., Evers, C., Stice, E., & Engels, R. C. M. E. (2012). The Eating Disorders Diagnostic Scale: Psychometric features within a clinical population and a cut-off point to differentiate clinical patients from healthy controls. *European Eating Disorders Review, 20,* 315–320.

Kraeplin, C. R. (2011). Minority females and the thin ideal: Ethnic versus mainstream fashion magazine and their effects on acculturation and body image in young Black and Latina women. *Journal of Research on Women and Gender, 2,* 50–82.

Kutlesic, V., Williamson, D. A., Gleaves, D. H., Barbin, J. M., & Murphy-Eberenz, K. P. (1998). The interview for the diagnosis of eating disorders-IVA: Application to DSM-IV diagnostic criteria. *Psychological Assessment, 10*(1), 41–48.

Luck, A., Morgan, J. F., Reid, F., O'Brien, A., Brunton, J., Price, C., Perry, L., & Ladey, J. H. (2002). The SCOFF questionnaire and clinical interview for eating disorders in general practice: Comparative study. *British Medical Journal, 325,* 755–756.

Maloney, M. J., McGuire, J. B., & Daniels, S. R. (1988). Reliability testing of a children's version of the eating attitude test. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*(5), 541–543.

Maloney, M. J., McGuire, J. B., Daniels, S. R., & Specker, B. (1989). Dieting behavior and eating attitudes in children. *Pediatrics*, *84*(3), 482–487.

Marcus, M., & Wing, R. R. (March 1983). Binge eating and dietary restraint. Proceedings of the Annual Meeting of the Society of Behavioral Medicine, Baltimore, Maryland, USA.

Marques, L., Alegria, M., Becker, A. E., Chen, C., Fang, A., Chosak, A., & Diniz, J. B. (2011). Comparative prevalence, correlates of impairment, and service utilization for eating disturbance across U. S. ethnic groups: Implications for reducing ethnic disparities in health care access for eating disorders. *International Journal of Eating Disorders, 44*(5), 412–420.

Mazzeo, S. E., Aggen, S. H., Andersen, C., Tozzi, F., Bulik, C. M. (2003). Investigating the structure of the Eating Inventory (Three-Factor Eating Questionnaire): A confirmatory approach. *International Journal of Eating Disorders, 34,* 255–264. doi:10.1002/eat.10180.

Mitchell, R. S., & Mazzeo, S. E. (2004). Binge eating and psychological distress in ethnically diverse undergraduates men and women. *Eating Behaviors, 5,* 157–169.

Mond, J. M., Myers, T. C., Crosby, R. D., Hay, P. J., Rodgers, B., Morgan, J. F., Lacey, H., & Mitchell, J. E. (2008). Screening for eating disorders in primary care: EDE-Q versus SCOFF. *Behaviour Research and Therapy, 46,* 612–622. doi:10.1016/j.brat.2008.02.003.

Morgan, J. F., Reid, F., Lacey, J. H. (1999). The SCOFF questionnaire: Assessment of a new screening tool for eating disorders. *British Medical Journal, 319*(7223), 1467–1468.

Marmorstein, N.R., von Ranson, K. M., Iacono, W. G., & Succop, P.A. (2007). Longitudinal associations between externalizing behaviors and dysfunctional eating attitudes and behaviors: A community-based study. *Journal of Clinical Child & Adolescent Psychology, 369*(1), 87–94. doi:10.1207/s15374424jccp3601_9.

Nangle, D. W., Johnson, W. G., Carr-Nangle, R. E., & Engler, L. B. (1994). Binge eating disorder and the proposed DSM-IV criteria: Psychometric analysis of the Questionnaire of Eating and Weight Patterns. *International Journal of Eating Disorders, 16*(2), 147–157.

Napolitano, M. A., & Himes, S. (2011). Race, weight, and correlates of binge eating in female college students. *Eating Behaviors, 12,* 29–36.

O'Neill, S. K. (2003). African American women and eating disturbances: A meta-analysis. *Journal of Black Psychology, 29*(1), 3–16.

Palmer, R., Christie, M., Cordle, C., Davies, D., & Kenrick, J. (1987). The Clinical Eating Disorder Rating Instrument (CEDRI): A preliminary description. *International Journal of Eating Disorders, 6,* 9–16.

Palmer, R., Robertson, D., Cain, M., & Black, S. (1996). The Clinical Eating Disorder Rating Instrument (CEDRI): A validation study. *European Eating Disorders Review, 4,* 149–156.

Perry, L., Morgan, J. F., Reid, F., Brunton, J., O'Brien, A., Luck, A., & Lacey, J. H. (2002). Screening for symptoms of eating disorders: Reliability of the SCOFF screening tool with written compared to oral delivery. *International Journal of Eating Disorders, 32*(4), 466–472.

Peterson, C. B. (2005). Conducting the diagnostic interview. In I. E. Mitchell, & C. B. Peterson (Eds.), *Assessment of eating disorders* (pp. 32–58). New York: Guilford.

Podar, I., & Allik, J. (2009). A cross-cultural comparison of the Eating Disorder Inventory. *International Journal of Eating Disorders, 42*(4), 346–355.

Roberts, A., Cash, T. F., Feingold, A., & Johnson, B. T. (2006). Are Black-White differences in females' body dissatisfaction decreasing? A meta-analytic review. *Journal of Consulting and Clinical Psychology, 4,* 1121–1131.

Robinson, P., & Andersen, A. (1985). Anorexia nervosa in American Blacks. *Journal of Psychiatric Research, 19*(2/3), 183–188.

Sabik, N. J., Cole, E. R., & Ward, L. M. (2010). Are all minority women equally buffered from negative body image? Intra-ethnic moderators of the buffering hypothesis. *Psychology of Women Quarterly, 34,* 139–151.

Sancho, C., Asorey, O., Arija, V., & Canals, J. (2005). Psychometric characteristics of the children's eating attitudes test in a Spanish sample. *European Eating Disorders Review, 13*(5), 338–343.

Silber, T. J. (1986). Anorexia nervosa in Blacks and Hispanics. *International Journal of Eating Disorders, 5*(1), 121–128.

Smolak, L., & Levine, M. P. (1994). Psychometric properties of the Children's Eating Attitudes Test. *International Journal of Eating Disorders, 16,* 275–282.

Spitzer, R. L., Devlin, M., Walsh, B.T., Hasin, D., Wing, R., Marcus, M., Stunkard, A., Wadden, T., Yanoski, S., Agras, S., Mitchell, J., & Nonas, C. (1992). Binge eating disorder: A multi-site field trial of the diagnostic criteria. *International Journal of Eating Disorders, 11,* 191–203.

Steinberg, E., Tanofsky-Kraff, M., Cohen, M.L., Elberg, J., Freedman, R. J., Semega-Janneh, M., Yanovski, S. Z., & Yanovski, J.A. (2004). Weight patterns in the assessment of children's eating- disordered behaviors. *International Journal of Eating Disorders, 36*(2), 183–194. doi:10.1002/eat.20022.

Stice, E., Fisher, M., & Martinez, E. (2004). Eating Disorder Diagnostic Scale: Additional evidence of reliability and validity. *Psychological Assessment, 16,* 60–71.

Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological Assessment, 12,* 123–131.

Stojek, M. M. K., & Fischer, S. (2013). Thinness expectancies and restraint. Black and White college women: A prospective study. *Eating Behaviors, 14,* 269–273. doi.org/10.1016/j.eatbeh.2013.03.005.

Striegel-Moore, R. H., Wilfley, D.E., Pike, K. M., Dohm, F.A., & Fairburn, C. G. (2000). Recurrent binge eating in Black American women. *Archives of Family Medicine, 9,* 83–97.

Stunkard, A., & Messick, S. (1985). The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *Journal of Psychosomatic Research, 29*(1), 71–83. doi.org/10.1016/0022-3999(85)90010-8.

Taylor, J. Y., Caldwell, C. H., Baser, R. E., Faison, N., & Jackson, J. S. (2007). Prevalence of eating disorders among Blacks in the National Survey of American Life. *International Journal of Eating Disorders, 40,* 510–S14.

Taylor, J. Y., Caldwell, C. H., Baser, R. E., Matusko, N., Faison, N., & Jackson, J. S. (2013). V Classifications and correlates of eating disorders among Blacks: Findings from the National Survey of American Life. *Journal of Health Care for the Poor and Underserved, 24,* 289–310.

Thelen, M., Farmer, J., Wonderlich, S., & Smith, M. (1991). A revision of the Bulimia Test: The BULIT–R. *Psychological Assessment, 3,* 119–124.

Timmerman, G. M. (1999). Binge Eating Scale: Further assessment of validity and reliability. *Journal of Applied Biobehavioral Research, 4*(1), 1–12.

Vander Wal, J. S., & Thomas, N. (2004). Predictors of body image dissatisfaction and disturbed eating attitudes and behaviors in African American and Hispanic girls. *Eating Behaviors, 5,* 291–301.

Vander Wal, J. S., Stein, R., & Blashill, A. J. (2011). The EDE-Q, BULIT-R, and BEDT as self-report measures of binge eating disorder. *Eating Behaviors, 12,* 267–271. doi:10.1016/j.eatbeh.2011.07.006.

von Ranson, K. M., Cassin, S. E., Bramfield, T. D., & Fung, T. S. (2007). Psychometric properties of the Minnesota Eating Behavior Survey in Canadian university women. *Canadian Journal of Behavioural Science, 39*(2), 151–159. doi:10.1037/cjbs2007001.

von Ranson, K. M., Klump, K. L., Iacono, W. G., & McGue, M. (2005). The Minnesota Eating Behavior Survey: A brief measure of disordered eating attitudes and behaviors. *Eating Behavior, 6,* 373–392. doi:10.1016/jeatbeh.2004.12.002.

Wade, T.D., Byrne, S., Bryant-Waugh, R. (2008). The eating disorder examination: Norms and construct validity with and young and middle adolescent girls. *International Journal of Eating Disorders, 41,* 551–558.

Watkins, B., Frampton, I., Lask, B., & Bryant-Waugh, R. (2005). Reliability and validity of the child version of the Eating Disorder Examination: A preliminary investigation. *International Journal of Eating Disorders, 38*(2), 183–187.

Webb, J. B., & Hardin, A. S. (2012). A preliminary evaluation of BMI status in moderating changes in body composition and eating behavior in ethnically-diverse first-year college women. *Eating Behaviors, 13*(4), 402–405.

Wildes, J. E., Emery, R. E., & Simones, A. D. (2001). The roles of ethnicity and culture in the development of eating disturbance and body dissatisfaction: A meta-analytic review. *Clinical Psychology Review, 21*(4), 521–551.

Wilfley, D. E., Schreiber, G. B., Pike, K. M., Striegel-Moore, R. H., Wright, D. J., & Rodin, J. (1996). Eating disturbance and body image: A comparison of a community sample of adult black and white women. *International Journal of Eating Disorders, 20*(4), 377–387.

Williamson, D. A. (1990). *Assessment of eating disorders: Obesity, anorexia, and bulimia nervosa*. New York: Pergamon.

Assessing for Personality Disorders in the African American Client

14

Allison Kalpakci and Carla Sharp

Introduction

The practice of assessing personality pathology has greatly improved since its inception. Though still imperfect, PD assessment continues to evolve through enhanced understanding of human behavior and the development of precise measurement tools founded in sound scientific research (Clark et al. 1997; Gabbard 2005; Widiger and Samuel 2005). Despite these advances, personality is a multifaceted construct, intricately tied to the myriad factors that underlie human nature, including family environment, genetics, and culture. Both the presentation and perception of personality are transactionally related to these factors, rendering the assessment of PD a highly complex endeavor. Though scientific rigor bolsters the validity of newer and revised assessment instruments, clear delineations of normal versus pathological remain elusive and highly subject to predominant cultural norms. Countless complex behaviors contribute to the presentation of PD, and the rater—not the assessment instrument—ultimately makes the diagnostic decisions. As such, the process of assessing an individual for PD may be particularly vulnerable to an evaluator's subjective judgment (Ziegenbein et al. 2008; Farmer 2000).

Perhaps one of the most potentially problematic issues in PD assessment is bias. Broadly speaking, biased PD measurement can occur at any point in the assessment process including within the testing instrument, during administration, and in interpretation of scores. Researchers and test developers have addressed bias through various means including outlining more explicit and encompassing PD criteria, designing tests founded in psychological theories of PD, and conducting studies to examine tests' psychometric properties. However, it is important to note that these modifications arise from European American cultural influences (Okazaki and Sue 2003). For example, the majority of commonly used PD assessment instruments have been written by European Americans. Personality theories refer to conceptualizations of PD in European American culture. And psychometric studies have been conducted with predominantly European American samples. So, despite these modifications, one cannot assume that traditional PD assessment is equivalent for non-European Americans. Moreover, there is a risk for misdiagnosis in these groups (Dana 2000).

Misdiagnosing PD in non-European American ethnic groups is a serious issue (Dana 2000). Overpathologizing occurs when an individual's personality traits thought to be deviant or impairing are, in fact, reflections of a unique cultural or ethnic background. Underpathologizing due to inaccurate interpretation or misunderstanding of

C. Sharp (✉) · A. Kalpakci
Department of Psychology, University of Houston,
126 Heyne Building, Houston, TX 77204, USA
e-mail: csharp2@central.uh.edu

A. Kalpakci
e-mail: ahkalpakci@gmail.com

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_14, © Springer Science+Business Media New York 2015

the respondent's culture is also problematic. In either case, biased PD assessment may preclude an accurate diagnosis, ultimately delaying or preventing intervention and resource allocation where truly needed. Given the long-standing, functionally impairing nature of PD, as well as the stigma and weighty decisions often associated with a PD diagnosis, precision in assessment is crucial (Widiger and Samuel 2005).

Though some have viewed the field of psychology as less than culturally sensitive (Hall 1997), there has been a recent call to consider cross-cultural factors in psychological assessment (American Psychiatric Association 2000; American Psychological Association 2003; Stuart 2004). Research within this context has taken the perspective that sociocultural factors are entrenched with conceptualizations of personality and likely influence the assessment practice. These studies have examined the appropriateness of use of extant PD assessment measures with specific cultural and ethnic minority groups. The ethnic minority group of most focused empirical attention is African Americans. Findings suggest that some PD assessment tools show promise as valid measures of PD in African Americans, but it is also clear that empirical literature in this area is lacking. To avoid biased assessment of African Americans, clinical and research evaluators would benefit from a comprehensive review of the major PD assessment instruments with empirical data backing their use in this ethnic minority group.

Against this background, this chapter will focus on the major PD assessment instruments that have been empirically evaluated for their valid use with African Americans. First, sociocultural issues unique to African Americans will be discussed, with particular attention paid to those issues that impact the presentation and perception of PD. Next, the concept of PD and classification will be explained along with an overview of assessment tools available, including self-report measures and semistructured interviews. Following a discussion of test bias evaluation methods, the major PD assessments with established psychometric studies in African American samples and those assessments that show promise for use with this group will be reviewed. This review will be the main focus of the chapter and will include major findings on the validity, reliability, bias, and utility of these instruments. Finally, recommendations regarding PD assessment with African Americans for both clinical and research fields will conclude the chapter. At this point, we also wish to point out that although there has been significant advances in the assessment of PD in children and adolescents over the last decade (Sharp et al. 2012b; Sharp and Kine 2008), the current chapter will focus exclusively on PD assessment in adults.

Overview of Issues Related to Mental Health in African Americans

African Americans represent 13.6 % of the population, the largest ethnic minority group in the USA (U.S. Census Bureau 2010). Over recent decades, the African American population in general is becoming more diverse. Black Americans vary in regards to social status and level of education. The social and economic standing of African Americans has improved such that a robust African American middle class has emerged (Sue and Sue 2008). But despite these changes, a number of social issues including low socioeconomic status (SES), racial discrimination, and health disparities continue to negatively impact the lives of African Americans (Mays et al. 2007).

Economically, African Americans have traditionally been disadvantaged with 24.7 % of this population living below the poverty line (U.S. Census Bureau 2010). Compared with European Americans, African Americans have a threefold increase in rates of poverty (Williams and Williams-Morris 2000). Poverty is tied to a number of social problems such as low educational attainment, health problems, family hardship, increased rates of crime, and low psychological functioning (McLoyd 1998). Moreover, the net worth of Caucasian Americans is ten times more than that of African Americans and they enjoy higher SES (U.S. Census Bureau 2010). Low SES is linked to many problems, including poor mental health. In fact, adults with the lowest

SES are nearly three times more likely to have a psychiatric condition than adults with the highest SES (Holzer et al. 1986; Hudson 2005; Robins and Regier 1991).

Though an apparent decrease in racial discrimination has been observed since 1960, discrimination against African Americans continues to be a serious issue in the USA. William and Williams-Morris (2000) provided data on stereotypes that European Americans commonly hold about African Americans. These included beliefs that black Americans were unintelligent, prone to violence, and prefer state-assisted funding like welfare to getting a job. Compared with stereotypes against other ethnic groups like Asian Americans and Hispanic Americans, the viewpoints against African Americans were decidedly more negative.

A study conducted by Kessler et al. (1999) found that 49 % of the African Americans in the sample experienced at least one instance of overt racial discrimination (e.g., racial profiling by police, receiving substandard services, etc.) and a surprisingly 81 % of the sample reported experiencing less overt forms of racial discrimination (e.g., subtle stereotyping). These negative attitudes affect hiring and lead to negative treatment of African Americans in the work place and in school (Neckerman and Kirschenman 1991). They also affect the general well-being of African Americans (Williams et al. 1997).

Another particularly debilitating effect of racism is its effect on psychological functioning (Sellers et al. 2006; Sellers and Shelton 2003; Williams et al. 2003). Studies have shown that chronic exposure to racial discrimination leads to stress and increased emotional reactivity (Sellers et al. 2003; Fang and Myers 2001). Other studies have found that exposure to racial discrimination is linked to higher levels of paranoia (Frueh et al. 1996) and depression, obsessive-compulsive disorder, and anxiety disorder (Kessler et al. 1999; Klonoff et al. 1999). It is clear that the issues that African Americans may face living in the USA can negatively impact psychological well-being and lead to serious mental health problems. Despite this elevated risk, there is generally lower usage of outpatient mental health care in African American populations compared with European Americans (Alegría et al. 2002; Harris et al. 2005).

A number of factors may contribute to these low consumption rates of mental health services. Some studies have found that African American clients reported feeling distrustful of mental health care personnel (Thompson et al. 2004) or that the psychological services received did not resonate with their needs, leaving them without effective solutions (Adebimpe 1984; Carter 1991). Muntaner and Parsons (1996) cited low insurance coverage as an additional hindrance. Some studies have found that there may be greater stigma associated with mental health problems in some African American communities compared with Caucasian communities which causes reluctance to seek psychiatric help (Link and Phelan 1999; Snowden 2001). Collectively, these factors suggest that African Americans with mental illness may not be receiving the services they need. For individuals whose mental illness is severe, this is a serious risk. One particularly debilitating class of psychiatric illness linked with especially low rates of mental health service consumption among African Americans is *PD* (Bender et al. 2007).

Personality Disorder

With the recent release of the newest Diagnostic and Statistical Manual in 2013 (DSM-5; APA 2013), experts within the fields of psychology and psychiatry continue to engage in a heated debate about how PDs should be classified. Initial proposals to revise the PD classification structure would have greatly altered the way PD is assessed, moving from a categorically based conceptualization of PD to a more dimensionally based taxonomy. However, the APA did not revise the DSM criteria according to the PD Workgroup's recommendations, and therefore PD continues to be classified categorically. It should be noted that the Workgroup's recommendations were included in Section III (conditions requiring further research). A more thorough discussion of the major issues surrounding this debate is beyond the scope of this chapter. (To seek additional information, see First 2005; Shedler et al. 2010; Skodol et al. 2011; Widiger and Samuel 2005.)

According to the DSM-IV-TR (APA 2000), a PD is a constellation of stable pathological personality traits and symptomatology that significantly impairs the inter- and intrapersonal functioning of an individual. The DSM-IV-TR classifies PDs into three types or clusters. Each cluster is grouped by the typical features that characterize the PD presentation. Cluster A includes PDs with odd or eccentric characteristics (Paranoid, Schizoid, and Schizotypal). Cluster B includes PDs with emotional, dramatic, or impulsive characteristics (Antisocial, Borderline, Histrionic, and Narcissistic). Cluster C includes PDs with fearful or anxious characteristics (Obsessive-Compulsive, Dependent, and Avoidant). A diagnosis of PD not otherwise specified (PD NOS) is given when an individual's personality functioning is significantly reduced, causes clinically significant impairment, but does not meet criteria for any of the ten specific PD types.

Aside from the requirement that PD traits are impairing, the DSM-IV-TR also indicates that they deviate significantly from societal norms. This requirement necessitates those norms first be defined and accepted (either implicitly or explicitly). In the USA, where the norms follow predominantly European American societal trends, this could be a considerable issue for non-European Americans or individuals raised in non-European American-based cultures. The DSM-IV-TR recognizes this issue and warns against diagnosing behaviors considered deviant according to predominant cultural norms, but that are native or specifically relevant to the individual's cultural origins (e.g., avoid diagnosing paranoid PD when behavior is evoked by perceived racial discrimination; APA 2000). However, the APA's suggestions on considering cultural and ethnic group differences in assessing PD do not suffice in ensuring valid PD assessment. Comprehensive and unbiased PD assessment relies heavily on the expertise of informed and culturally sensitive evaluators that use well-validated and reliable PD assessment instruments (Dana 2000).

Personality Disorder Assessment: Instruments and Bias

Despite the functionally impairing nature of PD, the systematic and comprehensive assessment of PD in clinical and research settings is not common practice (Michonski et al. 2012; Noblin et al. 2013; Sharp et al. 2012a). Instead, Axis I diagnoses are emphasized and PDs can be overlooked. Often times, however, an Axis I diagnosis does not fully explain an individual's symptomatology, underlying the need for additional assessment of relevant psychopathology. In addition to providing a more comprehensive view of an individual's psychological functioning, conducting an assessment of PD may help predict an individual's response to intervention, what obstacles might be expected throughout treatment, or additional variables that should be accounted for in research (Widiger and Samuel 2005). As such, it is highly recommended that PD be considered when conducting psychological assessment (Sharp et al. 2012a; Widiger and Samuel 2005).

There are various tools that clinicians and researchers use to measure PD. These tools fall under three broad categories: self-report inventories, interview-based schedules, and individual history as gathered from criminal and medical records and "important other"—rated (e.g., peer, partner, therapist, and parent) reports. In assessing African Americans for PD, it is generally recommended that all three sources of information inform diagnostic decisions (Widiger and Boyd 2009). However, the aim of this chapter is to provide information regarding the appropriateness of employing commonly used, empirically researched assessment instruments with African Americans, so only self-report inventories and interview-based measures will be discussed in the remainder of the chapter.

Self-Report Inventories

Self-report measures of PD follow a highly structured format. Items that describe PD traits and other psychopathology are rated by the respondent depending on their level of self applicability. There are several advantages to using self-report PD inventories over other measurement types (Widiger and Boyd 2009; Widiger and Chaynes 2003). Practically speaking, self-report inventories are relatively quick and inexpensive to administer, allowing the evaluator to assess a wide range of behaviors and symptomatology in a short amount of time (Noblin et al. 2013). Furthermore, self-report inventories are heavily researched, so their psychometric properties are better known relative to interview-based measures. But, there are also disadvantages to using self-report inventories. First, personality is a highly complex construct, with countless factors contributing to its presentation. A self-report inventory completed by a respondent does not incorporate other important sources of information like evaluator-rated behavioral observations, so the assessment scores may not fully capture the presentation of PD in that individual. Additionally, self-report measures often have higher rates of false positives due to high sensitivity and low specificity. This makes self-report inventories ideal for use as a first step in PD assessment, or as a PD screener (Chang et al. 2011; Morey 1991). If the inventory indicates elevated levels of any PD traits, an interview-based assessment should be employed (Noblin et al. 2013; Widiger and Samuel 2005).

Interview-Based Measures

There are two common types of interview-based PD assessments: semistructured and unstructured interviews. In clinical settings, unstructured interviews are often preferred, but due to their unsystematic nature, there is generally low diagnostic interrater reliability. Garb (2005) found that unstructured interviews increased the likelihood of culturally biased PD assessment. Furthermore, a number of studies have found that ethnic minorities are more likely to receive an inaccurate diagnosis when assessed using an unstructured interview (Widiger and Boyd 2009). Because of these reasons and because there is little to no research on unstructured interviews in African Americans, this chapter will focus on semistructured interviews.

Semistructured interviews are often used in research settings, but are recommended for use in clinical settings as well (Sharp et al. 2012; Widiger and Boyd 2009). The format of this type of assessment tool allows the evaluator to ask open-ended questions and include behavioral observations in diagnostic decisions, and ultimately the decision to endorse an item relies on the clinical expertise of the evaluator. These interviews also provide a more systematic way to assess for PD comprehensively, yielding better replicability and reliability across interviews than unstructured interviews (Widiger and Boyd 2009). However, interview-based measures are often more expensive and time-consuming to administer than self-report inventories. In addition, there is far less research regarding their psychometric properties. As such, it is recommended that interview-based measures are used in tandem with self-report measures.

Evaluating Bias in Personality Disorder Assessment Test Content

In recent decades, researchers have focused efforts on understanding how tests may fail to measure PD equivalently across ethnic and cultural groups. In measuring PD in African Americans, the goal of this line of research is to evaluate whether commonly used tests of PD are biased for use with this group. To this end, researchers have taken a few approaches to evaluate bias: methods based on ipso facto definitions of bias, regression model methods, and factor analytic and item response theory (IRT) approaches. Those in favor of taking an ipso facto approach have examined differences in PD assessment scale scores across African Americans and Caucasians. Inherent in this method is the assumption that any difference in test scores is indicative of bias. These studies typically match participants on education level and SES so that conclusions drawn about group differences are not rendered questionable due to potentially confounding variables.

However, this scale comparison procedure is considered insufficient according to many researchers who assert that the presence of differences does not imply bias (Cleary 1968). They argue that differences in group test scores might exist because of differences in measurement, and that these scores may not reflect a group's "real" construct scores. To test this type of bias, researchers use regression models to examine whether test scores predict differences in conceptually relevant, extra-test measures of personality functioning. One criticized feature of this method is the use of highly *Eurocentric* (or European–American-based) extra-test variables. According to this perspective, if the extra-test measure is also Eurocentric, then to claim one that is truly evaluating the criterion validity of PD assessment scales in African American samples is an untenable assumption. Instead, instruments developed based on conceptualizations of PD in African American culture should serve as these extra-test measures (Dana 2000).

The most promising method thus far for evaluation of bias is the use of IRT. The way IRT models the relation between observed item-responses and the latent trait is different from approaches employed in classical test theory (CTT; Sharp et al. 2006, 2012c). Instead of summarizing the psychometric properties of a scale with omnibus statistics (such as item-total correlations or Cronbach alpha), thereby averaging across levels of individual variation (Santor et al. 1994), IRT approaches model how the probability of responding to an item—here this is equivalent to endorsing a PD symptom—varies as a function of the location along a latent continuum or dimension of variation (Santor et al. 1994). IRT methods do not use summary statistics that apply to groups of individuals, such as correlations, but can define a model for the individual response patterns that comprise the raw data. Because item response patterns can be modeled directly within an IRT framework, no information in the data is lost. Crucially, for the purposes of cross-cultural assessment of PD, IRT, through differential item functioning (DIF) analyses, provides the opportunity to distinguish between bias at the level of the item (i.e., the item does not accurately probe for the symptom for a particular cultural group) and bias at the level of the latent trait (i.e., the disorder does not express itself through a particular symptom in a particular cultural group).

African American Personality Disorder Assessment: A Review of Empirical Studies

The following is a review of the major PD assessment instruments with empirical research examining their use with African Americans. Self-report measures including the Minnesota Multiphasic Personality Inventory (MMPI-2; Butcher et al. 1989), the Personality Assessment Inventory (PAI; Morey 1991), and the Millon Clinical Multiaxial Inventory (MCMI; Million 2006) will be reviewed. Interview-based measures including the Shedler and Westen Assessment Procedure (SWAP-II; Shedler and Westen 2007) and the Psychopathy Checklist-Revise (PCL-R; Hare 1991) will also be discussed. Other PDs that are often used in practice, but with no published data on African Americans will also be mentioned in brief. Following this review, recommendations for clinicians and researchers for assessment with African Americans with PD will be provided.

Self-Report Inventories

Minnesota Multiphasic Personality Inventory (MMPI-2)

The MMPI-2 (Butcher et al. 1989) is a 557-item assessment considered to be a gold standard test of personality and psychological functioning. Since its initial publication in 1943 (MMPI; Hathaway and McKinley 1943), the measure has undergone several revisions and additions. Changes that have significantly altered the test include the renorming and restandardization of the test population that resulted in the publication of the MMPI-2 in 1989 (Butcher et al. 1989), the adjustment of the psychometric flaws of the clinical scales that led to the addition of the restructured clinical scales in 2003 (Tellegen et al.

2003), and the release of the MMPI for use in adolescents (MMPI-A; Butcher et al. 1992). In addition, the restructured form (MMPI-2-RF; Ben-Porath and Tellegen 2008) was developed in 2008 and appears to be a more streamlined substitute of the MMPI-2; however, because this measure is relatively new and lacks empirical research on its use in African Americans, the focus of this section will center on the MMPI-2.

The MMPI-2 is appropriate for use with adults over 18 years of age with a reading level of at least grade 6. The 557 Items (rated as either *True* or *False*) assemble into several scales and dozens of subscales. The major scales of the MMPI-2 include the clinical and validity scales. The clinical scales measure common diagnoses and include Scale 1 (*HS*; Hypochondriasis), Scale 2 (*D*; Depression), Scale 3 (*Hy*; Hysteria), Scale 4 (*Pd*; Psychopathic Deviate), Scale 5 (*MF*; Masculinity/Femininity), Scale 6 (*Pa*; Paranoia), Scale 7 (*Pt*; Psychasthenia), Scale 8 (*Sc*; Schizophrenia), Scale 9 (*Ma*; Hypomania), and Scale 0 (*Si*; Social Introversion). The validity scales assess the level of exaggeration or overendorsement of psychological symptomology, the occurrence of item nonresponsiveness and the level of item response consistency, and the level of underendorsement or minimization of symptomatology. Additionally, there are many supplementary, content, and restructured clinical scales.

The original norming population of the MMPI (Hathaway & McKinley 1943) was composed of married individuals living in Minnesota, and it is highly likely that few (if any) minorities were included in the sample (Cox et al. 2009). Thus, the appropriateness of its use in non-European Americans has been questioned. Early studies addressed this issue by evaluating test bias of the MMPI for use in ethnic minorities, the most widely researched group being African Americans. Findings from that research indicated that African Americans and European Americans scored significantly different on some scales (Timbrook and Graham 1994). Though other researchers questioned the clinical meaningfulness of such differences (Greene 1987) as their pattern was somewhat inconsistent (Timbrook and Graham 1994), it was suggested that the MMPI may be biased for use in African Americans (Gynther and Green 1980). The newly standardized MMPI-2 (Butcher et al. 1989) attempted to address this bias by producing a more representative normative group than that of the original MMPI (Schinka and LaLone 1997).

A relatively large body of research has examined the appropriateness of use of the newer, restandardized MMPI-2 in African Americans (Arbisi et al. 2002; Castro et al. 2008; Dean et al. 2008; Frueh et al. 1996; Frueh et al. 1997; McNulty et al. 1997; Timbrook and Graham 1994; *for a review see* Hall et al. 1999). Across these studies, significant scale score mean differences and differential predictive validity between African Americans and European Americans have been observed in community and clinical samples. However, the differences are generally inconsistent, difficult to replicate, and often statistically, but not *clinically*, significant. In general, the consensus of this research and of experts in the field is that the MMPI is not a biased PD assessment instrument for use with African Americans (Cox et al. 2009). Nevertheless, the fact that differences (albeit small) exist suggests a need for careful assessment practice and additional research. A review of selected studies in community, clinical, and special populations is presented below.

The MMPI-2 for Use in African Americans from Community Populations Timbrook and Graham (1994) analyzed differences on the MMPI-2 clinical and validity scales in a subset of the European Americans and African Americans from the original MMPI-2 restandardization norming population (Butcher et al. 1989). They conducted two studies. The first study examined mean scale differences between African Americans and European Americans matched on family income, years of education, and age. Results from this study revealed significant differences between these two groups, with African American men scoring significantly higher on Scale 8 (Schizophrenia) than European American men. Additionally, African American women scored significantly higher on Scales 4 (Psychasthenia), 5 (Masculinity/Femininity), and 9 (Hypomania) than did European American women. Despite these findings, mean score differences were minimal at less than 5 *T*-score points, calling into question the clinical meaningfulness of the findings.

The second (Timbrook and Graham 1994) study served to contextualize the findings from Study 1, by examining the differential predictive validity of five (2, 4, 7, 9, and 0) of the MMPI-2 scales by testing whether there were differences in how well the MMPI-2 scales predicted participant partner-rated psychological functioning variables across the two groups (the same sample in Study 1). Results revealed that Scale 7 significantly underpredicted partner-rated anxiety scores for African American women compared to their European American counterparts. However, no differences were reported in the predictive ability of the remaining four MMPI-2 clinical scales. The findings from this study are notable in that they counter previous thought that the MMPI-2 scales overpathologize African Americans. However, because Timbrook and Graham (1994) did not examine differential predictive validity in the remaining clinical scales or in any of the other MMPI-2 scales, these findings were somewhat limited in their conclusive scope. Furthermore, given that a community sample was used in these studies, whether the MMPI-2 was biased for use in African Americans from clinical populations was indeterminable.

Waller et al. (2000) used IRT to examine measurement bias on MMPI items and scales. They found that 38 % of the items on each scale were, on average, biased against Blacks or Caucasians, but that item bias canceled out when aggregated at the scale score level, and therefore concluded that DIF does not lead to differential test functioning.

The MMPI-2 for Use in African Americans from Clinical Populations Two studies addressed the question of appropriateness of use of the MMPI-2 in clinical African American populations (McNulty et al. 1997; Arbisi et al. 2002). McNulty et al. (1997) extended Timbrook and Graham's (1994) research and examined whether the MMPI-2 clinical and content scales differentially predicted conceptually similar therapist-rated measures of personality symptomatology across outpatient European and African Americans. Results demonstrated no significant differential correlations between the MMPI-2 scales and the therapist-rated variables across the two groups with the exception of Scale 9 (Hypomania). African American women scored significantly higher on mean Scale 9 score than did European American women, yet therapist-rated measure of agitation did not reflect this same difference. This finding may suggest that Scale 9 overpathologizes African American women in this domain of psychopathology.

Arbisi et al. (2002) further extended MMPI-2 research in clinical populations by analyzing mean scale differences between African Americans and European Americans in an inpatient psychiatric setting. Results from this study demonstrated that some of the MMPI-2 scales (2, 4, 8, 9, DEP, ASP, AAS, BIZ, and MAC-R) differentially predicted extra-test variables of psychological functioning (e.g., Axis II diagnosis, bipolar diagnosis, antidepressant use, anxiety diagnosis, psychosis, etc.) in European American and African American participants. The majority of these differences indicated that the scales underpredicted psychopathology (albeit with small effect sizes) in African American participants, a finding that contradicted McNulty et al.'s (1997) results.

In a more recent study, Castro et al. (2008) examined the differential predictive accuracy of MMPI-2 clinical and restructured clinical (RC) scales in European American and African American in a community outpatient sample. Arguing that use of therapist or partner-rated non-MMPI-2 indices of symptomology for comparison with MMPI-2 scale scores may have impacted results in other studies of this kind (Arbisi et al. 2002; McNulty et al. 1997; Timbrook and Graham 1994), Castro et al. used client self-reported extra-test measures of psychopathology. This study followed similar procedures as the previous studies mentioned and found that African American clients scored higher on Scales 1, RC1 (Somatic Complaints), RC3 (Cynicism), RC6 (Ideas of Persecution), and RC8 (Aberrant Experiences), but these difference did not accurately predict the non-MMPI self-report scores, suggesting these scales overpathologize African American clients.

Summary of MMPI-2 Studies with African Americans Overall, the findings from the aforementioned studies reveal an inconsistent pattern of MMPI-2 scale score differences across African Americans and Caucasians. Some studies showed that a few MMPI-2 scales underpathologized African Americans (Arbisi et al. 2002; Castro et al. 2008; Timbrook and Graham 1994), while others demonstrated the opposite pattern (McNulty et al. 1997). As such, research reveals no systematic, clinically significant pattern of group differences in MMPI-2 scale scores. Given these findings, the majority of researchers and PD assessment experts alike have concluded that the MMPI-2 is a valid and unbiased instrument for use with African American clients (Arbisi et al. 2002; Castro et al. 2008; Cox et al. 2009; McNulty et al. 1997; Timbrook and Graham 1994), countering Gynther's (1972) suggestion for unique interpretive guidelines for African Americans.

Taken together, these mixed findings illuminate the importance of examining standard of practice when using and interpreting the MMPI-2 with African Americans. Despite a study that found that the ethnicity of the client (African American vs. Caucasian) did not impact clinician interpretation of MMPI-2 clinical scales, there remains a risk of a client's ethnicity influencing clinical judgment (Ozaki & Sue (1995); Luepnitz et al. 1982). Dahlstrom et al. (1986) asserted that taking the individual into account when interpreting test results is necessary. Velasquez et al. (2000) elaborated this assertion by warning against heavy reliance on computerized interpretive reports when evaluating ethnic minorities. Such reports do not consider cultural factors, and thus should be used with caution in these cases. Additionally, they suggest using all MMPI-2 scales (e.g., Content, Supplementary, and Restructured Clinical) in interpretation as the clinical and validity scales alone may not fully capture the PD of African American clients. As such, use of abbreviated version of the MMPI-2 (first 370 items of MMPI-2, includes Validity and Clinical scales only) with African Americans should be avoided. These mixed findings also propose a need for more research in this area. Hill et al. (2012) called for the development of measures tailored for the specific ethnic group of interest that would serve as more culturally relevant extra-test correlates (Table 14.1).

The Personality Assessment Inventory (PAI)

The personality assessment inventory (PAI; Morey) is a 344-item, scale-based questionnaire that measures psychological functioning across a wide range of domains. The PAI has 22 non-overlapping scales that can be classified into four domains: (1) Validity Scales which measure respondent approach to the test (e.g., *inconsistency*, *infrequency*, *negative impression*, and *positive impression*), (2) Clinical Scales which assess psychiatric diagnostic categories (e.g., *somatic complaints*, *anxiety*, *anxiety-related disorder*, *depression*, *mania*, *paranoia*, *schizophrenia*, *borderline features*, *antisocial features*, *alcohol problems*, and *drug problems*), (3) Treatment Scales which assess factors related to treatment of clinical disorders (e.g., *aggression*, *suicidal ideation*, *stress*, *nonsupport*, and *treatment rejection*), and (4) Interpersonal Scales that assess interpersonal functioning (e.g., *dominance* and *warmth*). Respondents rate items on a scale from 1–4 (1—false, 2—somewhat true, 3—mainly true, 4—very true). The PAI is appropriate for use in adults aged 18 and over with at least a fourth grade reading level. Additionally, the PAI-A (Morey 2007) was developed for use with adolescents aged 12–18 years.

The PAI is a viable alternative to the MMPI-2 for a number of reasons. In terms of practicality of usage, it is a shorter test, taking around 50 min to complete. In addition, it includes subscales that capture personality features within the respective PD (borderline and antisocial). For example, the borderline scale assesses the identity problems, negative relationships, self harm, and affective instability features prominent in the disorder (APA 2000; Zanarini 2003). Though specific combinations of other scales can be used to diagnose other PDs, no empirical studies have been conducted to evaluate the validity of this procedure.

Table 14.1 Self-report inventories and interview-based measures for the assessment of personality pathology in African American adults

Assessment name	Disorders assessed	Relevant research findings	Recommendations for use with African Americans
Minnesota Multiphasic Personality Inventory-2 (MMPI-2)	Personality and personality disorder; psychological functioning	Overall, no evidence of systematic pattern of bias for use with African Americans. Findings of ethnic group differences are inconsistent, and sizes of differences found were of small effect (Arbisi et al. 2002; Castro et al. 2008; McNulty et al. 1997; Timbrook and Graham 1994)	Use of this assessment with African Americans is generally recommended Do not rely solely on Clinical and Validity scales for interpretation; rather, incorporate Content, Supplementary, and Restructured Clinical scales. Avoid heavy reliance on computerized interpretative reports
Personality Assessment Inventory (PAI)	Personality disorder and psychopathology	Few studies have examined the validity of using this assessment with African Americans. Findings reveal small differences between African American and Caucasian scale scores, except for PAR scale (Morey 2007)	Use separate norms to convert raw scores to *T*-score for African Americans (provided in Appendices C & D of PAI manual; Morey 2007)
Millon Clinical Multiaxial Inventory (MCMI-III)	Personality and personality disorder, psychiatric symptoms	Research on use of this assessment with African Americans is lacking. However, studies examining the use of the MCMI-II have found that African Americans have elevated scores relative to Caucasian Americans (Choca et al. 1990; Hamberger and Hastings 1992)	Little research regarding this instrument exists; thus, do not solely rely on this instrument when assessing for PD in African Americans
Shedler–Westin assessment procedure	Personality and personality disorder	Results from unpublished study support the convergent and discriminant validity of factors (Westen et al. 2007)	May serve as a promising assessment in the future for clinical settings, but more research needed before widespread use with African Americans
Hare Psychopathy Checklist Revised (PCL-R)	Antisocial personality disorder and psychopathology	Mixed findings regarding ethnic group differences. Some studies have found different patterns of scores and underlying factor structures across African Americans and Caucasians (Kosson et al. 1990). One study found potential bias of the antisocial facet of the PCL-R (Sullivan et al. 2006). Studies with African American women yield support for validity of this measure, though cut-off score of 30 may be too high for female inmates (Vitale et al. 2002)	Use of PCL-R in African American inmates is generally recommended, though evaluators should recognize risk of bias from relying heavily on criminal records for assessment; rather, personality traits should be emphasized

The psychometric properties of the PAI are strong and empirically supported (Morey 1991). In developing this measure, the authors took a comprehensive, multiphasic, statistic-, and theory-based approach (Holden 2000; Morey 1991). Stringent quantitative methods were applied to selection of items and thorough evaluations of bias and appropriateness of items and constructs were conducted. For example, a bias review panel examined items that would potentially overpathologize behaviors that are normal within certain subcultures. Psychometric studies support internal consistency and test–retest reliability and that the PAI scales appropriately converge with relevant scales of the MMPI-2 (Morey 2007). The PAI has been normed for various populations including both normal and clinical populations as well as university students (Morey 1991); however, few African American respondents were included in these samples (less than 6 % in the normative sample, 2.8 % in the college sample, and 12.5 % in the clinical sample).

In the PAI Professional Manual, Morey (2007) provided information regarding differences in PAI scores due to ethnicity. Findings from this research suggest that differences are small for most of the scales except for PAR (Paranoia) in which African Americans may score seven *T*-scores higher than Caucasians. Provided in the manual is a table (p. 94, Table 14.1) with estimates of variance accounted for by ethnicity in each of the clinical scales. Though differences across ethnic groups were minimal, Morey suggests using separate norms to convert raw scores to *T*-scores for African Americans (provided in Appendices C and D).

To our knowledge, only one peer-reviewed study has explicitly evaluated the use of the PAI in African Americans. Atlerman et al. (1995) examined reliability and concurrent validity of the PAI in African American and Latino American methadone maintenance treatment patients. Results revealed that the psychometric properties of the PAI in this sample were mostly consistent with the original standardization sample. Few differences were observed between men and women participants, and Latino participants reported more symptomatology than did African Americans. Results also support the current validity of the PAI and extra-test measures of psychiatric diagnoses and symptomology. However, it was revealed that 30 % of the profiles were invalid, thus questioning the use of the validity scales with these populations.

Given that only one peer-reviewed study has analyzed the use of the PAI in an African American sample, it is crucial that additional research on the appropriateness of use of this instrument in African Americans is conducted. The fact that this instrument is relatively new may be a potential reason for the lack of empirical attention. There are many gaps to be filled. Studies should be conducted in African Americans from a wide variety of samples, including healthy and clinical. These studies should examine the reliability, validity, and potential bias in this measure by comparing scale scores between African Americans and other groups, and examining the predictive accuracy of those scores to extra-test measures of conceptually relevant variables.

For both clinical and research applications, interpretation of the PAI with African Americans should be carefully considered and other sources of information and knowledge about the respondent should be incorporated into any diagnostic decision. Though the psychometric properties of this instrument are robust for the standardization sample, whether this strength carries over to use in African Americans or other ethnic minorities is unknown.

The Millon Clinical Multiaxial Inventory (MCMI-III)

The MCMI-III (Millon 2006) is a self-report instrument that assesses the DSM-IV-TR PDs (as well as Axis I disorders). The 175 items are rated as either *True* or *False* and take approximately 25 minutes to complete. The test is appropriate for adults 18 years and older with at least an eighth grade reading level. There are three types of scales included in the MCMI-III: PD Scales (e.g., *avoidant, borderline, schizotypal, antisocial, narcissistic, histrionic, dependent, aggressive, compulsive, passive-aggressive, self-*

defeating, paranoid, and *depressive*), Axis I Scales (e.g., *major depression, dysthymic disorder, bipolar: manic disorder, anxiety disorder, posttraumatic stress disorder, alcohol dependent, drug dependence, somatoform disorder, thought disorder,* and *delusional disorder*) and Validity Scales (e.g., *debasement, validity, disclosure,* and *desirability*).

The norming sample used for the original MCMI was proportionate to minority group representations in the USA and separate standard score conversions are available for African Americans, as well as for men, women, and Hispanics in the professional manual. Although this test was designed for use in clinical populations, norms exist for nonclinical populations. Despite these efforts for greater generalizability, research with the original MCMI (1987; aligned with the DSM-III) revealed that African American scores on this test were often elevated relative to Caucasian scores. For example, Choca et al. (1990) found that African Americans scored significantly different from Caucasian Americans on both item and scale scores. In addition, the scale scores differentially predicted DSM-III diagnoses for Caucasians and African Americans. Hamberger and Hastings (1992) matched outpatients on age, education, and employment status and found that African Americans scored significantly higher than Caucasians on delusional disorder, paranoid, and narcissistic scales.

To our knowledge, only one study has examined ethnic differences in MCMI-III scores. Ghafoori and Hierholzer (2010) recently examined the relationship between race, ethnicity, and PDs in a sample of African American, Hispanic, and Caucasian combat veterans. Results revealed that Hispanic veterans were more likely to exhibit Cluster A personality pathology that non-Hispanic veterans. No other differences were observed across the groups.

Additional clarifying research is needed to evaluate appropriateness of use of the MCMI with African Americans. However, until published findings are available, it is generally suggested that clinicians and researchers working with African Americas use other PD assessments to complement use of this measure (Choca et al. 1992).

Other Self-Report Inventories

There are other measures of PD that are commonly used in practice and in research settings. These measures include self-report inventories like the Personality Diagnostic Questionnaire-4th edition (PDQ-4+; Hyler 1994), a 99-item measure that assesses for DSM-IV-TR PDs. Though this is a widely used measure in clinical practice due to its relative brevity, there is no research regarding the validity of its use with African Americans. Additionally, the Wisconsin Personality Inventory (WISPI; Klein 1985), the Multisource Assessment of Personality Pathology (MAPP; Clifton et al. 2007), and the Coolidge Axis II Inventory (CATI; Coolidge 1993) are also used, but with no research regarding validity of use with African Americans, other more well-validated assessments should likely be used. (For further information regarding the advantages and disadvantages of these assessments, see Widiger and Boyd 2009.)

Further, members of the DSM-5 PD workgroup developed a maladaptive personality trait model and corresponding self-report instrument called the Personality Inventory for DSM-5 (Krueger et al. 2012). It is essential that this measure, alongside other self-report measures of the normal variation in personality traits (e.g., the NEO Personality Inventory-Revised (NEO-PI-R; Costa and McCrae 1992) are investigated for invariance across African American and other groups.

Interview-Based Measures

Shedler Westen Assessment Procedure (SWAP-II)

The SWAP-II (Shedler and Westen 2007) is a 200-item PD assessment designed to balance the external validity of its use in clinical practice while still maintaining strong internal validity. Assessors using the SWAP-II may either employ a structured interview or spend a minimum of 6 clinical hours with a client before scoring the SWAP items (i.e., descriptions of personality features across a wide range of domains). This instrument employs a Q-sort procedure such that the assessor sorts each item into one of eight categories (0—*not descriptive* to 7—*highly descriptive*) based on its descriptive accuracy of the interviewee's personality. There is a fixed number of items that must be assigned per category so that across clients, the distribution of scores is consistent, minimizing error variance due to evaluator individual rating tendencies. Researchers and clinicians complete the assessment on a Microsoft Excel program that utilizes scoring algorithms to produce a SWAP profile. This profile consists of scales and PD *T*-scores as well as probable DSM-IV PD diagnoses. The PD scores indicate client similarity to the PD prototype. Additionally, a personality health index (PHI) indicates level of personality health. There is a SWAP available for use with adolescents (SWAP-II-A; Westen et al. 2003).

At this time, the SWAP is not a widely used method to assess PDs in clinical settings (Widiger and Boyd 2005); however, it was designed for optimal clinical applicability and thus it may likely serve as a promising assessment procedure in the future. A number of studies have examined the reliability and validity of this method; however, no peer-reviewed studies have examined the appropriateness of use of the SWAP for African Americans. One unpublished study examined the construct validity of the SWAP using a multitrait–multimethod matrix approach developed by Campbell and Fiske (1959) in a sample of nonclinical, mostly African Americans from inner city regions. Results supported the convergent and discriminant validity of the factors (Westen et al. 2007). This study was part of a larger, ongoing study that will hopefully produce more research of this kind (Shedler and Westen 2007). In the meantime, it is important that future research focuses on examining the validity of using this PD assessment method with African Americans.

The Psychopathy Checklist-Revised (PCL-R)

The PCL-R (Hare 1991) is a 20-item interview-based measure considered to be the gold standard assessment for adult psychopathy. Used primarily in forensic settings, the PCL-R assesses both DSM-IV-TR antisocial PD, as well as four additional psychopathic features including glib charm, lacking in empathy, shallow affect, and arrogance. Though the PCL-R is interview based, the assessor also incorporates information from the respondent's medical, clinical, and criminal record in completing the assessment. Because of this integrative process, completion of the PCL-R can be labor intensive and requires a significant amount of time (Widiger and Boyd 2009). A dimensional score is provided with cut-off indicated at 30, a number that has been supported in studies with male offenders (Hare 2011).

Studies that have examined the validity of the PCL-R generally support its psychometric properties (Hare et al. 1999); however, these studies have mostly used Caucasian samples and thus the validity of its use with African Americans cannot be automatically assumed. There are a few studies that have examined the validity of the use of the PCL-R in African Americans. Kosson et al. (1990) found different patterns of correlations between personality scores and psychopathy for Caucasians and African Americans inmates. Furthermore, factor analysis revealed differential underlying factor structures in the two groups. Cooke et al. (2001) looked at the PCL-R in Caucasian and African American individuals and found that though the instrument itself was not biased (similar underlying factor structure and similar item functioning for both groups), biased *use* of the instrument is still possible.

In a more recent study, Sullivan et al. (2006) examined the PCL-R in Caucasian, Latino, and African American male prison inmates and found that there were minor differences in PCL-R scores in African Americans compared to the other two groups. These differences, however, did not differentially predict extra-test variables across the groups. In terms of the antisocial facet of the PCL-R, there was evidence of a significantly stronger relationship between antisocial personality disorder (ASPD) and IQ and weaker relationship between ASPD and other psychopathological variables in African Americans relative to the other groups, suggesting a potential issue with antisocial construct in African Americans. Despite findings of problems with the ASPD construct, the authors conclude that the PCL-R is a valid assessment for psychopathy in African Americans.

Two studies have investigated the PCL-R in incarcerated African American women. Vitale et al. (2002) examined the validity of the PCL-R in a large sample of Caucasian and African American inmates. Results from this study supported convergent and discriminant validity of the instrument in female inmates, though it was observed that the cut-off score of 30 may be too high for female inmates. This may be due to lower base rates of psychopathy in women or it could also be an artifact of the paucity of studies of female psychopaths with large sample sizes. Overall, however, the study yielded support for the reliability and validity of the PCL-R in women. In a more recent study, Sturek et al. (2008) examined the use of the structured clinical interview for DSM-IV (SCID-II) as a screener for the PCL-R in a sample of female inmates (56% African American) and concluded that using the conduct disorder scale as part of a standard mental health intake interview would reliability select those individuals that required further assessment with the PCL-R.

Taken together, the research on the PCL-R suggests that the PCL-R may be a valid instrument to assess psychopathy in African American male and female inmates. The cut-off score most used is 30, but may be too high a threshold for females. Research should examine this issue and help determine a more appropriate cut-off score for women. Vitale et al. (2002) suggested factor analytic and confirmative analyses and IRT analyses to better understand the item functioning of the PCL-R in females.

Additionally, some think that the heavy reliance on details gleaned from criminal records to complete the PCL-R may lead to an overinflation of ASPD or psychopathy in prison populations and have suggested that PD traits should instead be more heavily incorporated in PCL-R to avoid this bias (Skeem and Cooke 2010; Widiger and Boyd 2009). Clinicians should acknowledge this as a risk for bias and mitigate its effects in practice.

Other Interview-Based Personality Disorder Measures

In addition to the measures listed above, there are also a number of interview-based measures based on the DSM-IV-TR PD criteria. These include the PD inventory (PDI-IV; Widiger et al. 1995), the structure clinical interview for DSM-IV Axis II PDs (SCID-II; First and Gibbon 1997), the diagnostic interview for personality disorders (DIPD-IV; Zanarini et al. 1987), the structured interview for DSM-IV PDs (SIDP-IV; Pfohl et al. 1997), the international PD examination (IPDE; Loranger et al. 1997), and the inventory of interpersonal problems (IIP; Horowitz et al. 2000). Additional interview-based measures that assess for specific PDs, include the revised diagnostic interview for borderlines (DIB-R; Zanarini et al. 1989) and the diagnostic interview for narcissism (DIN; Gunderson et al. 1990). Though these measures are in wide use in clinical settings, the validity of their use with African Americans is unclear. Research should examine the appropriateness of these interview schedules for use with African Americans before they are commonly used with this ethnic group. (For further information regarding the advantages and disadvantages of these assessments, see Widiger and Boyd 2009.)

Recommendations

Clinicians can increase their likelihood of conducting unbiased PD assessment with African Americans in a number of ways (Morris 2000). First, clinicians must increase their knowledge about African American sociocultural issues. This can be accomplished by attending multicultural awareness trainings and workshops, consulting with multicultural counseling expert, and voicing concerns about multicultural insensitivities and/or bias to supervisors. If necessary, referring a client to an evaluator more familiar with African American culture is also recommended (APA 2000). Second, clinicians should remain cognizant of implicit racial relationship dynamics between evaluator and respondent by increasing mindfulness of one's own reactions, misconceptions, and assumptions throughout the assessment process. One suggestion is to keep a written record of one's perceptions about the client throughout the assessment process, especially those perceptions that are potentially ethnoculturally biased. Along this line, it is crucial that the client feels that he/she can provide feedback regarding the level of cultural sensitivity he/she experienced within the assessment process. Maintaining rapport through the assessment process will help facilitate this dialogue.

Third, it is the duty of a clinician to stay abreast of current multicultural PD assessment research and recommended procedures. In regards to a specific assessment PD protocol (Widiger and Samuel 2005), it is recommended to first administer a well-validated self-report inventory, next administer an interview-based measure, and then gather information about the client from partners, therapists, friends, or family members. Taking this multimethod approach creates a comprehensive assessment practice and understanding of the individual, decreasing the likelihood for bias. In regards to choices of specific tests and scales, it is recommended to use tests with empirical data backing their use with African Americans and when using other less-validated tests, ensure this is not the sole source of information guiding a diagnostic decision. Further, avoid using computerized reports and include additional scales where possible. By incorporating all of these recommendations, clinicians can move toward greater multicultural competence. Achieving this competence is an evolving process—one that a clinician must proactively pursue throughout his or her career.

In some ways, researchers have an even greater responsibility in fostering a multiculturally competent PD assessment practice with African Americans. This is perhaps due to the fact that relatively little research has been conducted in this area and there are many possible future directions. First, studies must be conducted to better understand the specific role ethnicity plays in PD. In particular, this research should focus on specific factors relevant to African Americans like racism, discrimination, and poverty (Lindsey and Cuellár 2000). Conducting qualitative research to further explore these issues might be useful (Lee and Ramirez 2000). Second, studies should be conducted to understand the evaluator–client relationships and how evaluator bias influences PD assessment with African Americans. Third, researchers should develop and evaluate tests that measure acculturation, racial identity attitudes, and the ethnocentricity of worldview for African Americans. There are some measures already available (Helms and Parham 1996; Landrine and Klonoff 1994; Kelsey and Ransom 1996), but they are not in widespread use currently and require further research for use with PD assessment instruments. Fourth, additional research examining PD assessment instrument bias for use with African Americans is needed. In determining whether test scores under- or over-pathologize African Americans, more culturally relevant extra-test measures must be developed and validated. Given that research traditionally focuses on examining bias in self-report inventories, studies should begin to focus on examining bias in interview-based measures as well. The use of IRT and DIF in this regard is essential. Fifth, it is essential that the DSM-5 Section III recommendations be incorporated in future research agendas as it is unlikely that the current DSM-IV conceptualization of PD will be indefinitely retained. Understanding how dimensional traits of maladaptive personality functioning vary as a function of race at the level of mean scores, traits and items will do much in moving forward the PD field. Finally, studies should take a holistic approach by comparing African American PD scores across entire tests (i.e., PAI vs. MMPI-2, MMPI-2 vs. SCID-II, etc.) so that we can begin to establish the relative validity of PD assessment tests.

Conclusion

The overall aim of this chapter was to review the major PD assessment tools available for use with African Americans. The chapter began with an overview of sociocultural issues relevant to Africans, especially those that influence the perception and presentation of PD. Then, the PD construct was discussed, along with an overview of common PD assessment practices. Next, the issue of test bias and its evaluation were introduced. This segued into a review of the major PD assessment instruments that have been evaluated for their appropriateness of use with Africans Americans. Finally, recommendations for clinicians and researchers were provided.

It should now be clear that the assessment of PD is a complex process vulnerable to cross-cultural bias. When evaluating PD in Africans Americans for both clinical and research purposes, one must proactively aim to mitigate the risks of biased assessment. This can be accomplished by choosing well-validated instruments, taking a multimethod assessment approach, and examining and improving one's standard of multicultural assessment practice. Given the implications, decisions, and prognosis often associated with PD diagnosis, it is critical that PD assessment is conducted in a systematic and unbiased manner across all ethnic and cultural groups.

References

Adebimpe, V. R. (1984). American Blacks and psychiatry. *Transcultural Psychiatric Research, 21,* 83–111.

Alegría, M., Canino, G., Ríos, R., Vera, M., Calderón, J., Rusch, D., & Ortega, A. N. (2002). Mental health care for Latinos: Inequalities in use of specialty mental health services among Latinos, African Americans, and non-Latino Whites. *Psychiatric Services, 53*(12), 1547–1555.

Alterman, A. I., Zaballero, A. R., Lin, M. M., Siddiqui, N., Brown, L. S., Rutherford, M. J., & McDermott, P. A. (1995). Personality Assessment Inventory (PAI) scores of lower-socioeconomic African American and Latino methadone maintenance patients. *Assessment, 2*(1), 91–100.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., rev.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.) Washington, DC: American Psychiatric Association.

American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. *American Psychologist, 58*(5), 377–402.

Arbisi, P. A., Ben-Porath, Y. S., & McNulty, J. (2002). A comparison of MMPI-2 validity in African American and Caucasian psychiatric inpatients. *Psychological assessment, 14*(1), 3–15.

Bender, D. S., Skodol, A. E., Dyck, I. R., Markowitz, J. C., Shea, M. T., Yen, S., Sanislow, C. A., Pinto, A., Zanarini, M. C., McGlashan, T. H., Gunderson, J. G., Daversa, M. T., & Grilo, C. M. (2007). Ethnicity and mental health treatment utilization by patients with personality disorders. *Journal of consulting and clinical psychology, 75*(6), 992–999.

Ben-Porath, Y. S., & Tellegen, A. (2008). *Minnesota Multiphasic Personality Inventory-2 restructured form: MMPI-2-RF*. Minneapolis: University of Minnesota Press (NCS Pearson, Inc).

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.

Butcher, J. N., Williams, C. L., Graham, J. R., Archer, R., Tellegen, A., Ben-Porath, Y. S., et al. (1992). *MMPI-A manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105.

Carter, R. T. (1991). Cultural values: A review of empirical research and implications for counseling. *Journal of Counseling and Development, 70,* 164–173.

Castro, Y., Gordon, K. H., Brown, J. S., Anestis, J. C., & Joiner, T. E. (2008). Examination of racial differences on the MMPI-2 clinical and restructured clinical scales in an outpatient sample. *Assessment, 15*(3), 277–286.

Chang, B., Sharp, C., & Ha, C. (2011). The criterion validity of the borderline personality feature scale for children in an adolescent inpatient setting. *Journal of Personality Disorders, 25*(4), 492–503.

Choca, J. P., Shanley, L. A., Peterson, C. A., & Van Denburg, E. (1990). Racial bias and the MCMI. *Journal of Personality Assessment, 54,* 479–490.

Choca, J. P., Shanley, L. A., & Van Denburg, E. (1992). *Interpretive guide to the Millon Clinical Multiaxial Inventory*. Washington, DC: American Psychological Association.

Clark, L. A., Livesley, W. J., & Morey, L. (1997). Personality disorder assessment: The challenge of construct validity (special feature). *Journal of Personality Disorders, 11*(3), 205–231.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement, 5*(2), 115–124.

Clifton, A., Turkheimer, E., & Oltmanns, T. F. (2007). Improving assessment of personality disorder traits through social network analysis. *Journal of Personality, 75*(5), 1007–1032.

Cooke, D. F., Kosson, D. S., & Michie, C. (2001). Psychopathy and ethnicity: Structural, item, and test generalizability of the psychopathy checklist-revised (PCL-R) in Caucasian and African American participants. *Psychological Assessment, 13,* 531–542.

Coolidge, F. L. (1993). *The Coolidge Axis II inventory: Manual*. Clermont: Synergistic Office Solutions.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa: Psychological Assessment Resources.

Cox, C. A., Weed, N. C., & Butcher, J. N. (2009). The MMPI-2: History, interpretation and clinical Issues. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (pp. 250–276). New York: Oxford University Press.

Dahlstrom, W. G. (1986). Ethnic status and personality measurement. In W. G. Dahlstrom, D. Lacher, & L. E. Dahlstrom, *MMPI Patterns of American Minorities* (pp. 3–23). Minneapolis: University of Minnesota Press.

Dana, R. H. (2000). *Handbook of cross-cultural and multi-cultural personality assessment*. Mahwah: Lawrence Erlbaum Associates.

Dean, A. C., Boone, K. B., Kim, M. S., Curiel, A. R., Martin, D. J., Victor, T. L., Zeller, M. A., & Lang, Y. K. (2008). Examination of the impact of ethnicity on the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) fake bad scale. *The Clinical Neuropsychologist, 22*(6), 1054–1060.

Fang, C. Y., & Myers, H. F. (2001). The effects of racial stressors and hostility on cardiovascular reactivity in African American and Caucasian men. *Health Psychology, 20*(1), 64–70.

Farmer, R. (2000). Issues in the assessment and conceptualization of personality disorders. *Clinical Psychology Review, 20,* 823–851.

First, M. B. (2005). Clinical utility: A prerequisite for the adoption of a dimensional approach in DSM. *Journal of abnormal psychology, 114*(4), 560–564.

First, M. B., & Gibbon, M. (1997). *User's guide for the structured clinical interview for DSM-IV axis II personality disorders: SCID-II*. Washington, DC: American Psychiatric Press.

First, M. B., & Gibbon, M. (2004). The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II).

Frueh, B. C., Smith, D. W., & Libet, J. M. (1996). Racial differences on psychological measures in combat veterans seeking treatment for PTSD. *Journal of Personality Assessment, 66*(1), 41–53.

Frueh, B. C., Gold, P. B., & de Arellano, M. A. (1997). Symptom overreporting in combat veterans evaluated for PTSD: Differentiation on the basis of compensation seeking status. *Journal of Personality Assessment, 68*(2), 369–384.

Gabbard, G. O. (2005). Mind, brain, and personality disorders. *American Journal of Psychiatry, 162*(4), 648–655.

Garb, H. N. (2005). Clinical Judgment and Decision Making. *Clinical Psychology, 1*(1), 67–89.

Ghafoori, B., & Hierholzer, R. W. (2010). Personality patterns among Black, White, and Hispanic combat veterans. *Psychological Trauma: Theory, Research, Practice, and Policy, 2*(1), 12.

Greene, R. L. (1987). Ethnicity and MMPI performance: A review. *Journal of Consulting and Clinical Psychology, 55,* 497–512.

Gunderson, J., Ronningstam, E., & Bodkin, A. (1990). The diagnostic interview for narcissistic patients. Archives of General Psychiatry, 47, 676–680.

Gynther, M. D. (1972). White norms and black MMPIs: A prescription for discrimination?. *Psychological Bulletin, 78*(5), 386–402.

Gynther, M. D., & Greene, S. B. (1980). Accuracy may make a difference, but does a difference make for accuracy? A response to Prtichard and Rosenblatt. *Journal of Consulting and Clinical Psychology, 48,* 268–272.

Hall, C. C. I. (1997). Cultural malpractice: The growing obsolescence of psychology with the changing US population. *American Psychologist, 52*(6), 642–651.

Hall, G. C. N., Bansal, A., & Lopez, I. R. (1999). Ethnicity and psychopathology: A meta analytic review of 31 years of comparative MMPI/MMPI-2 research. *Psychological Assessment, 11*(2), 186–197.

Hamberger, L. K., & Hastings, J. E. (1992). Racial differences on the MCMI in an outpatient clinical sample. *Journal of Personality Assessment, 58*(1), 90–95.

Hare, R. D. (1991). *The revised psychopathy checklist*. Toronto: Multi-Health Systems.

Hare, R. D. (2011). The Hare PCL-R: Some issues concerning its use and misuse. *Legal and Criminological Psychology, 3*(1), 99–119.

Hare, R. D., Cooke, D. J., & Hart, S. D. (1999). Psychopathy and sadistic personality disorder. In T. Millon, P. H. Blaney, & R. D. Davies (Eds.), *Oxford textbook of psychopathology* (pp. 555–584). London: Oxford University Press.

Harris, K. M., Edlund, M. J., & Larson, S. (2005). Racial and ethnic differences in the mental health problems and use of mental health care. *Medical Care, 43*(8), 775–784.

Hathaway, S. H. & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory*. Minneapolis: The University of Minnesota Press.

Helms, J. E., & Parham, T. A. (1996). The racial identity attitude scale. *Handbook of tests and measurements for Black populations, 2,* 167–174.

Hill, J. S., Robbins, R. R., & Pace, T. M. (2012). Cultural validity of the Minnesota Multiphasic Personality Inventory-2 empirical correlates: Is this the best we can do? *Multicultural Counseling and Development, 40,* 104–116.

Holden, R. R. (2000). Are there promising MMPI substitutes for assessing psychopathology and personality? Review and prospect. In R. H. Dana (Ed.), *Handbook of cross-cultural and multicultural personality assessment* (pp. 267–292). Mahwah: Laurence Erlbaum Associates.

Holzer, C. E., Shea, B. M., Swanson, J. W., & Leaf, P. J. (1986). The increased risk for specific psychiatric disorders among persons of low socioeconomic status. *American Journal of Social Psychiatry, 6,* 259–271.

Horowitz, L. M., Strauß, B., & Kordy, H. (2000).*Inventory of Interpersonal Problems (IIP)*. Göttingen: Hogrefe

Hudson, C. G. (2005). Socioeconomic status and mental illness: Tests of the social causation and selection hypotheses. *American Journal of Orthopsychiatry, 75*(1), 3–18.

Hyler, S. E. (1994). *Personality questionnaire PDQ-4+*. New York: State Psychiatric Institute.

Kelsey, R. C., & Ransom, R. M. (1996). A comparison of African and European groups utilizing a worldview opinion questionnaire. In R. L. Jones (Ed.), *Handbook of tests and measurements for the black population*. Hampton: Cobb & Henry.

Kessler, R. C., Mickelson, K. D., & Williams, D. R. (1999). The prevalence, distribution, and mental health correlates of perceived discrimination in the United States. *Journal of health and social behavior, 40,* 208–230.

Klein, M. (1985). *Wisconsin Personality Inventory (WISPI)*. Madison: University of Wisconsin.

Klonoff, E. A., Landrine, H., & Ullman, J. B. (1999). Racial discrimination and psychiatric symptoms among Blacks. *Cultural Diversity and Ethnic Minority Psychology, 5*(4), 329–339.

Kosson, D. S., Smith, S. S., & Newman, J. P. (1990). Evaluating the construct validity of psychopathy in black and white male inmates: three preliminary studies. *Journal of Abnormal Psychology, 99*(3), 250–259.

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*(9), 1879–1890. doi:10.1017/S0033291711002674.

Landrine, H., & Klonoff, E. A. (1994). The African American acculturation scale: Development, reliability, and validity. *Journal of Black Psychology, 20*(2), 104–127.

Lee, R. M., & Ramirez, M., III. (2000). The history, current status, and future of multicultural psychotherapy. In I. Cuellár & F. A. Paniagua (Eds.), *Handbook of multicultural mental health: Assessment and treatment of diverse populations* (pp. 280–303). San Diego: Academic Press.

Lindsey, M. L., & Cuellár, I. (2000). Mental health assessment and treatment of African Americans: A multicultural perspective. In I. Cuellár & F. A. Paniagua (Eds.), *Handbook of multicultural mental health: Assessment and treatment of diverse populations* (pp. 195–207). San Diego: Academic Press.

Link, B. G., & Phelan, J. C. (1999). Public conceptions of mental illness: Labels, causes, dangerousness and social distance. *American Journal of Public Health, 89,* 1328–1333.

Loranger, A. W., Janca, A., & Sartorius, N. (Eds.). (1997). *Assessment and diagnosis of personality disorders: The ICD-10 international personality disorder examination (IPDE)*. Melbourne: Cambridge University Press.

Luepnitz, R. R., Randolph, D. L., & Gutsch, K. U. (1982). Race and socioeconomic status as confounding variables in the accurate diagnosis of alcoholism. *Journal of Clinical Psychology, 38*(3), 665–669.

Mays, V. M., Cochran, S. D., & Barnes, N. W. (2007). Race, race-based discrimination, and health outcomes among African Americans. *Annual Review of Psychology, 58,* 201–225.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53*(2), 185–204.

McNulty, J. L., Graham, J. R., Ben-Porath, Y. S., & Stein, L. A. R. (1997). Comparative validity Of MMPI-2 scores of African American and Caucasian mental health center clients. *Psychological Assessment, 9*(4), 464–470.

Michonski, J. D., Sharp, C., Steinberg, L., & Zanarini, M. C. (2012). An item response theory analysis of the DSM-IV borderline personality disorder criteria in a population-based sample of 11- to 12-year-old children. *Personal Disorders: Theory, Research, and Treatment, 4*(1), 15–22.

Millon, T. (2006). *Millon Clinical Multiaxial Inventory-III manual* (3rd ed.). Minneapolis, MN: NCS Pearson.

Morey, L. C. (1991). *Personality assessment inventory professional manual*. Odessa: Psychological Assessment Resources.

Morey, L. C. (2007). *Personality assessment inventory-adolescent: Professional manual*. Odessa: Psychological Assessment Resources.

Morris, E. F. (2000). Assessment practices with African Americans: Combining standard assessment measures within an Africentric orientation. In R. H. Dana (Ed.), *Handbook of cross-cultural and multicultural personality assessment* (pp. 573–603). Mahwah: Lawrence Erlbaum Associates, Inc.

Muntaner, C., & Parsons, P. E. (1996). Income, stratification, class and health insurance. *International Journal of Health Services, 26,* 655–671.

Neckerman, K. M., & Kirschenman, J. (1991). Hiring strategies, racial bias, and inner-city workers. *Social Problems, 38,* 433–437.

Neighbors, H. W., Trierweiler, S. J., Munday, C., Thompson, E. E., Jackson, J. S., Binion, V. J., & Gomez, J. (1999). Psychiatric diagnosis of African Americans: diagnostic divergence in clinician-structured and semistructured interviewing conditions. *Journal of the National Medical Association, 91*(11), 601–612.

Noblin, J. L., Venta, A., & Sharp, C. (2013). The Validity of the MSI-BPD among inpatient adolescents. *Assessment, 21*(2), 210–217. doi:10.1177/1073191112473177.

Okazaki, S., & Sue, S. (2003). Methodological issues in assessment research with ethnic minorities. In A. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (3rd ed., pp. 349–367). Washington: American Psychological Association.

Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured interview for DSM-IV personality: SIDP-IV*. Washington, DC: American Psychiatric Press, Inc.

Robins, L. N., & Regier, D. A. (1991). *Psychiatric disorders in America: The epidemiologic catchment area study*. New York: Free Press.

Santor, D. A., Ramsay, J. O., & Zuroff, D. C. (1994). Nonparametric item analyses of the beck depression inventory: Evaluating gender item bias and response option weights. *Psychological Assessment, 6*(3), 255–270.

Schinka, J. A., & LaLone, L. (1997). MMPI-2 norms: Comparisons with a census-matched subsample. *Psychological Assessment, 9,* 307–311.

Sellers, R. M., & Shelton, J. N. (2003). The role of racial identity in perceived racial discrimination. *Journal of Personality and Social Psychology, 84*(5), 1079–1092.

Sellers, R. M., Caldwell, C. H., Schmeelk-Cone, K. H., & Zimmerman, M. A. (2003). Racial identity, racial discrimination, perceived stress, and psychological distress among African American young adults. *Journal of Health and Social Behavior, 44,* 302–317.

Sellers, R. M., Copeland–Linder, N., Martin, P. P., & Lewis, R. H. (2006). Racial identity matters: The relationship between racial discrimination and psychological functioning in African American adolescents. *Journal of Research on Adolescence, 16*(2), 187–216.

Sharp, C., & Kine, S. (2008). The assessment of juvenile psychopathy: Strengths and weaknesses of currently used questionnaire measures. *Child and Adolescent Mental Health, 13*(2), 85–95.

Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The short mood and feelings questionnaire (SMFQ): A unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. *Journal of Abnormal Child Psychology, 34*(3), 379–391.

Sharp, C., Green, K., Venta, A., Pettit, J., & Zanarini, M. C. (2012a). The incremental validity of borderline personality disorder relative to major depressive disorder for suicidal ideation and deliberate self-harm in adolescents. *Journal of Personality Disorders, 26*(6), 927–938.

Sharp, C., Ha, C., Michonski, J., Venta, A., & Carbone, C. (2012b). Borderline personality disorder in adolescents: Evidence in support of the childhood interview for DSM-IV borderline personality disorder in a sample of adolescent inpatients. *Comprehensive Psychiatry, 53*(6), 765–774.

Sharp, C., Steinberg, L., Yaroslavsky, I., Hofmeyr, A., Dellis, A., Ross, D., & Kincaid, H. (2012c). An item response theory analysis of the problem gambling severity index. *Assessment, 19*(2), 167–175. doi:10.1177/1073191111418296.

Shedler, J., Beck, A., Fonagy, P., Gabbard, G. O., Gunderson, J., Kernberg, O., Michels, R., & Westen, D. (2010). Personality disorders in DSM-5. *American Journal of Psychiatry, 167*(9), 1026–1028.

Shedler, J., & Westen, D. (2007). The Shedler–Westen assessment procedure (SWAP): making personality diagnosis clinically meaningful. *Journal of Personality Assessment, 89*(1), 41–55.

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment, 22*(2), 433–445.

Skodol, A. E., Bender, D. S., Morey, L. C., Clark, L. A., Oldham, J. M., Alarcon, R. D., Krueger, R. F., Verheul, R., Bell, C. C., & Siever, L. J. (2011). Personality disorder types proposed for DSM-5. *Journal of Personality Disorders, 25*(2), 136–169.

Snowden, L. R. (2001). Barriers to effective mental health services for African Americans. *Mental Health Services Research, 3*(4), 181–187.

Stuart, R. B. (2004). Twelve practical suggestions for achieving multicultural competence. *Professional Psychology Research and Practice, 35*(1), 3–9.

Sue, D. W., & Sue, D. (2008). *Counseling the culturally diverse: Theory and practice*. New York: Wiley.

Sturek, J. C., Loper, A. B., & Warren, J. I. (2008). Psychopathy in female inmates: The SCID-II Personality Questionnaire and the PCL-R. *Psychological Services, 5*(4), 309–319.

Sullivan, E. A., Abramowitz, C. S., Lopez, M., & Kosson, D. S. (2006). Reliability and construct validity of the psychopathy checklist-revised for Latino, European American, and African American male inmates. *Psychological Assessment, 18*(4), 382–392.

Tellegen, A., Ben-Porath, Y. S., McNulty, J., Arbisi, P., Graham, J. R., & Kaemmer, B. (2003). *The MMPI-2 Restructured Clinical (RC) scales: Development, validation, and interpretation*. Minneapolis: University of Minnesota Press.

Thompson, V. L. S., Bazile, A., & Akbar, M. (2004). African Americans' perceptions of psychotherapy and psychotherapists. *Professional Psychology: Research and Practice, 35*(1), 19–26.

Timbrook, R. E., & Graham, J. R. (1994). Ethnic differences on the MMPI-2? *Psychological Assessment, 6,* 212–217.

Velasquez, R. J., Ayala, G. X., Mendoza, S., Nezami, E., Castillo-Canez, I., Pace, T., Choney, S. K., Gomez, F.C., Jr., & Miles, L. E. (2000). Culturally competent use of the Minnesota Multiphasic Personality Inventory-2 with minorities. In I. Cuellar & F. A. Paniagua (Eds.), *Handbook of multicultural mental health: Assessment and treatment of diverse populations* (pp. 390–412). San Diego: Academic Press.

Vitale, J. E., Smith, S. S., Brinkley, C. A., & Newman, J. P. (2002). The reliability and validity of the psychopathy checklist-revised in a sample of female offenders. *Criminal justice and behavior, 29*(2), 202–231.

Waller, N. G., Thompson, J. S., & Wenk, E. (2000). Using IRT to separate measurement bias from true group differences on homogeneous and heterogeneous scales: An illustration with the MMPI. *Psychological Methods, 5,* 125–146.

Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality diagnoses in adolescence: DSM-IV axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry, 160*(5), 952–966.

Westen, D., Waller, N., Blagov, P., Shedler, J., & Bradley, R. (2007). Measuring pathological personality traits by clinician-reports using the SWAP-II: Factor structure, validity, and retest reliability. Unpublished manuscript.

Widiger, T. A., & Boyd, S. E. (2009). Personality disorders assessment instruments. In J. N. Butcher (Ed.), *Oxford handbook of personality assessment* (pp. 336–363). New York: Oxford University Press.

Widiger, T. A., & Chaynes, K. (2003). Current issues in the assessment of personality disorders. *Current Psychiatry Reports, 5*(1), 28–35.

Widiger, T. A., Mangine, S., Corbitt, E. M., Ellis, C. G., & Thomas, G. V. (1995). *Personality Disorder Interview-IV. A semistructured interview for the assessment of personality disorders*. Professional manual. Odessa, FL: Psychological Assessment Resources.

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the diagnostic and statistical manual of mental disorders. *Journal of Abnormal Psychology, 114*(4), 494–504.

Williams, D. R., & Williams-Morris, R. (2000). Racism and mental health: The African American experience. *Ethnicity and Health, 5*(3–4), 243–268.

Williams, D. R., Yu, Y., Jackson, J. S., & Anderson, N. B. (1997). Racial differences in physical and mental health socio-economic status, stress and discrimination. *Journal of Health Psychology, 2*(3), 335–351.

Williams, D. R., Neighbors, H. W., & Jackson, J. S. (2003). Racial/ethnic discrimination and health: Findings from community studies. *American Journal of Public Health, 93,* 200–208.

Zanarini, M. C. (2003). Zanarini rating scale for borderline personality disorder (ZAN-BPD): A continuous measure of DSM-IV borderline psychopathology. *Journal of Personality Disorders, 17*(3), 233–242.

Zanarini, M. C., Frankenburg, F. R., Chauncey, D. L., & Gunderson, J. G. (1987). The diagnostic interview for personality disorders: Interrater and test-retest reliability. *Comprehensive Psychiatry, 28*(6), 467–480.

Zanarini, M. C., Gunderson, J. G., Frankenburg, F. R., & Chauncey, D. L. (1989). The revised diagnostic interview for borderlines: discriminating BPD from other axis II disorders. *Journal of Personality Disorders, 3*(1), 10–18.

Ziegenbein, M., Calliess, I. T., Sieberer, M., & Machleidt, W. (2008). Personality disorders in a cross-cultural perspective: Impact of culture and migration on diagnosis and etiological aspects. *Current Psychiatry Reviews, 4*(1), 39–47.

Screening for Cognitive Decline and Assessment of Dementia with African Americans

15

Lorraine T. Benuto, Cyndy G. Soto and Brian D. Leany

Assessment of Dementia with African Americans: A Neuropsychological Perspective

In 1998 the American Psychological Association (APA) published guidelines for the evaluation of dementia and age-related cognitive decline. These guidelines were updated in 2011 and currently include components on competence, ethical considerations, and procedural guidelines for conducting evaluations of dementia and age-related cognitive change. There are two general guidelines for psychologists related to competence: (1) psychologists are familiar with the prevailing diagnostic nomenclature and specific diagnostic criteria and (2) psychologists gain specialized competence in assessment and intervention with older adults. There are three general guidelines for ethical considerations: (1) psychologists are aware of the special issues surrounding informed consent in cognitively compromised populations; (2) psychologists seek and provide appropriate consultation, and (3) psychologists are aware of cultural perspectives and of personal and societal biases and engage in nondiscriminatory practice (APA 2012).

In addition to the above guidelines, there are eight procedural guidelines for conducting evaluations of dementia and age-related cognitive change (APA 2012).

The first guideline is that psychologists strive to obtain all appropriate information for conducting an evaluation including pertinent medical history and communicating with relevant health care providers. The second guideline is that psychologists conduct a clinical interview as part of the evaluation, in addition to objective neuropsychological testing. Key information that should be obtained in the clinical interview includes: onset and course of changes in cognitive functioning, pre-existing disabilities, educational and cultural background that could affect testing variability, general medical and psychiatric history, past neurologic history including prior head injuries or other central nervous system insults (strokes, tumors, infections, etc.), current psychiatric symptoms and significant life stressors, current prescription and over-the counter medication use, current and past use and abuse of alcohol and drugs, and family history of dementia. The remaining guidelines are that (3) psychologists are aware that standardized psychological and neuropsychological tests are important tools;

L. T. Benuto (✉) · C. G. Soto
Administrative Faculty, Department of Psychology
MS0296 University of Nevada, Reno, NV 89557, USA
e-mail: dr.benuto@gmail.com

L. T. Benuto · B. D. Leany
Northcentral University, Core Faculty in the Graduate School, 10000 E University Drive, Prescott Valley, AZ 86314, USA

C. G. Soto
e-mail: cyndy.g.soto@gmail.com

B. D. Leany
Lake's Crossing Center, Maximum Security Facility for Forensic Mental Health Services, 500 Galletti Way, Sparks, NV 89504, USA
e-mail: leanyb@gmail.com

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_15,

(4) when evaluating for cognitive and behavioral changes in individuals, psychologists attempt to estimate premorbid abilities; (5) psychologists are sensitive to the limitations and sources of variability and error in psychometric performance, and to the sources of error in diagnostic decision making; (6) psychologists make appropriate use of longitudinal data; (7) psychologists recognize that providing constructive feedback, support, and education as well as maintaining a therapeutic alliance, can be important parts of the evaluation process; (8) psychologists appropriately recommend interventions available to persons with cognitive impairment and their caregivers; and (9) psychologists are aware that full evaluation of possible dementia is an interdisciplinary, holistic process involving other health care providers. Guideline 9 also maintains that psychologists respect other professional perspectives and approaches, communicate fully, and refer appropriately to support integration of the full range of information for informing decisions about diagnosis, level of severity, and elements of the treatment plan (APA 2012).

The primary focus of this chapter is the APA's call for cultural considerations in the diagnosis of dementia. This coupled with the large number of diverse individuals living in the USA and the variation among ethnic groups in terms of prevalence rates of dementia necessitates a discussion on cultural factors that might impact accurate diagnoses and evaluation procedures. Thus, the focus of this chapter is on cultural considerations that must be made when assessing for cognitive decline with African Americans.

Indeed, accurate assessment of cognitive impairment with African Americans is both relevant and important as African Americans compose a large part of the aging population in the USA and represent a disproportionately large prevalence rate for dementia. According to the US Census (2013) in 2011 African Americans constituted 13.1 % of the US population and the African American older population was 3.2 million in 2008 (Administration on Aging 2010). Life expectancy at birth for African Americans (73.3 years) lags behind that of White Americans by approximately 5 years and African Americans have a higher prevalence of physiological conditions that increase their risk of developing dementia including heart disease, stroke, and diabetes (Hargrave 2010). In terms of prevalence rates of dementia, African Americans have a higher prevalence of vascular dementia (Froehlich et al. 2001) and Alzheimer's disease than White Americans (Alzheimer's Association n.d.).

In providing culturally sensitive services to African American older adults, it is important to possess background knowledge of the historical events that have shaped their culturally derived attitudes toward health and healthcare providers (for a detailed account of such historical events, see Hargrave 2010). Elderly African Americans may be suspicious of and reluctant to trust healthcare providers because of experiences with prejudice and discrimination by healthcare providers. This mistrust may be exacerbated by the fact that misdiagnosis and clinical bias continue to occur frequently in the psychiatric assessment and development of treatment plans for African American patients of all ages (Hargrave 2010). Thus, an important aspect of providing culturally sensitive services to African American older adults entails establishing trust and building rapport.

Specific to the assessment of cognitive decline a review of the extant literature revealed that African American elders tend to perceive dementia-induced cognitive decline as a normal part of healthy aging (Clark et al. 2005; Jett 2006). While certainly elders in general experience normal age-related cognitive decline, dementia does not fall under this "normal" rubric of cognitive decline. Complicating things further is a stigma African Americans have concerning dementia (Innes 2009). Thus routine screening in healthcare settings is one means by which all elders can be assessed to determine the nature and extent of any cognitive decline the elder may experience. Below, a brief summary of two major studies involving neuropsychological assessment of older African Americans follows as well as a review of screening measures for cognitive impairment and a review of measures specific to dementia is provided.

While this chapter is focused on providing a neuropsychological perspective for the assessment of cognitive decline two caveats are in order. First, the screening measures described herein can be used to screen for cognitive impairment that either is or is not specific to dementia. Second, while measures that are specific to the assessment of dementia are covered in this chapter, it may be appropriate for measures discussed in the neuropsychological assessment chapter of this text as well (e.g., if a diagnosis of dementia is made, it may be beneficial to identify specific areas of sparing and impairment to inform supportive and treatment interventions). As such the bulk of this chapter is focused on dementia-specific assessment measures and measures used to screen for cognitive decline. First, however, it is important to discuss two relevant large-scale studies.

Mayo's Older African Americans Normative Studies (MOANS)

Mayo's Older African Americans Normative Studies (MOANS) were a series of studies that aimed to generate age-appropriate norms for elderly African Americans on an array of commonly used neuropsychological assessment measures. Drawing on a sample of 309 community dwelling African Americans aged 56–94 years old residing in Jacksonville, Florida, MOANS researchers developed new norms for various neuropsychological measures in an effort to enhance diagnostic accuracy. Age corrected norms accounting for years of education were developed for: Mattis Dementia Rating Scale (DRS), Wechsler Adult Intelligence Scale-Revised (WAIS-R), Boston Naming Test, Controlled Oral Word Association, Category Fluency, Animal Naming, Token Test, Wide Range Achievement Test-3 (WRAT-3) Reading Subtest, Trail Making Test, Stroop Test, Judgment of Line Orientation, the Wechsler Memory Scale-Revised, and the Rey Auditory Verbal Learning Test (Rilling et al. 2005; Lucas et al. 2005a, b, c, d; Ferman et al. 2005). Thus, when using the above measures as part of a battery to assess for cognitive impairment, cognitive decline, or dementia we recommend the adjusted norms described above be used.

Baltimore Study of Black Aging (BSBA)

The Baltimore Study of Black Aging (BSBA) was a large-scale longitudinal study with a sample of 602 community dwelling African American older adults aged 50–95 years old. The goal of the BSBA was to obtain information about cognition, health, and psychosocial factors related to the population as well as the relationships between these. However, the BSBA did not generate data directly related to the amendment of screening measures for cognitive impairment (i.e., generating new norms). Nonetheless, from this study we can deduce that the subjective assessment of memory does not predict actual memory performance in older African Americans (Sims et al. 2008) and education is a more universal predictor of variability in late life cognition than self-reported cardiovascular health among older African Americans (Aiken-Morgan et al. 2010).

Assessing for Cognitive Impairment

Mini-Mental State Exam (MMSE) The Mini-Mental State Exam (MMSE; Folstein et al. 1975) is an 11-item screen, takes 5–10 min to administer, can easily be administered in primary care settings, and assesses ten areas of cognitive functioning including: orientation, registration, attention and calculation, recall, language, repetition, 3-stage command, reading, writing, and copying. The MMSE is typically used to evaluate cognitive impairment in individuals suspected of having dementia as well as to assess decreases in cognitive impairment in persons who have already been diagnosed with dementia. Specific to African Americans, the MMSE has been documented to have low specificity, yielding up to a 42% false–positive rate for cognitive impairment with this population compared to a 6% false–positive rate among Whites (Chin et al. 2011). Even when demographic factors such as age and education

(Hargrave 2006; Pedraza et al. 2012) are controlled for, cultural bias often remains (Chin et al. 2011). This bias may exist because of the historical unequal educational opportunities of African Americans. Specifically African American elders may have received lesser quality education, which could impact literacy levels; thus literacy level should be considered when interpreting MMSE scores more so than simply number of years of education (Schneider and Lichtenberg 2011). In fact, researchers have documented the explanatory power of reading ability (i.e., literacy levels) in predicting higher performance on the MMSE and attenuating the association between race and MMSE scores (Griffith et al. 2006; Crowe et al. 2008; Pedraza et al. 2012; Morgan et al. 2008). Noting the limited normative data for African-American older adults, Strickland et al. (2005) provided age and education stratified normative data for the MMSE based on a sample of 93 community dwelling African American elders aged 60 and over. More recently Hawkins et al. (2011) provided age and education stratified normative data for the MMSE from a sample of 298 community dwelling African American adults aged 55–87. This normative data distinguishes between MMSE total scores incorporating serial 7 s subtraction and MMSE total scores incorporating "world" spelled backward, suggesting that the serial 7 s item is a significantly more difficult test than a total MMSE score incorporating the word "world" spelled backward and that the two indicate a different metric (Hawkins et al. 2011). Indeed Pedraza et al. (2012) found that adjusting for age and quality of education does boost classification accuracy of dementia in elderly African American adults. Despite these findings, they deemed unadjusted MMSE scores as having satisfactorily high classification accuracy for clinical purposes. In addition to adjusting for age and quality of education, monitoring changes in MMSE scores in lieu of utilizing absolute scores has been proposed as an alternative in controlling for the low-specificity of the measure (Chin et al. 2011). Thus, with MMSE it is important that the clinician keep in mind the high false positive rate of cognitive decline that has been noted for this population. Moreover, literacy level adjustments should be assessed and considered in the interpretation of MMSE scores. Where cognitive decline is suspected, further assessment will be necessary including using collateral contacts. Lastly, the MMSE may be useful in terms of establishing baseline cognitive performance and may be used to track cognitive decline over time.

Montreal Cognitive Assessment (MoCA) The Montreal Cognitive Assessment (MoCA; Nasreddine et al. 2005) is a screen for mild cognitive impairment that takes approximately 10 min to administer. It assesses multiple cognitive domains including: visuospatial abilities, executive function, short-term memory, attention, concentration, working memory, language, and orientation to time and place. The MoCA is especially useful as a screening tool in persons with MMSE scores of 26 or higher as it is more sensitive to impairments in executive functioning than the MMSE (Dong et al. 2010). The MoCA is typically used to assess for subtle cognitive deficits and has normative data from a large (N = 2653), ethnically diverse (African Americans constituted 52 % of the sample), cardiac-diseased, population-based sample. While differences in MoCA scores by ethnic groups were not compared it is important to note that African Americans have not been excluded from research on the MoCA. Nevertheless, researchers have indicated that caution is needed when interpreting MoCA scores in multicultural situations, especially among groups with lower education (Rossetti et al. 2011).

In a sample of 94 cognitively healthy (Clinical Dementia Rating = 0) elderly persons aged 60–90 in Bangkok, Thailand researchers (Julayanont et al. 2013) examined the effect of education and literacy on performance on the MoCA. The researchers found education effects in tasks assessing memory, executive, and visuospatial function: abstraction, repetition, trail-B, cube copy, clock drawing test, letter fluency, delayed recall and total MoCA scores. Additionally, the researchers found literacy effects, independent of education, on attention and working memory tasks and language related tasks: digit forwards, digit backwards, naming, vigilance test, serial-7 subtraction and total orientation score and its sub-items. These findings suggest that caution is needed when interpreting MoCA scores in groups with lower education as well as those with lower literacy levels. In fact, the authors recommend the development of a suitable cognitive screening test for illiterate and low educated elderly (Julayanont et al. 2013).

Cognistat The Cognistat was formerly known as the Neurobehavioral Cognitive Status Examination (NCSE; Kiernan et al. 1987). It consists of a graded series of questions (each section begins with a "screen" question and subsequent items are only administered if the individual passes the item) that assess the following: orientation; attention language; constructional ability; memory; calculation; and verbal reasoning. Unlike the MMSE, which only produces a single score, the Cognistat allows for an evaluation of various domains and purportedly has greater specificity than the MMSE (Osato et al. 1993, 1994). Specific to this chapter, researchers found that among African American elders (age 60+) there were not strong education and age effects (Strickland et al. 2005). It is worth noting that while Strickland et al. only collected data from African Americans, they did note that relative to findings identified with Euro American samples (e.g., Kiernan et al. 1987) African Americans performed poorer in the areas of attention, repetition, naming, construction, memory, calculations, and judgment were noted (Strickland et al. 2005). Given these differences in level of performances compared to the standardization group, Strickland et al. (2005) provided age and education stratified normative data from their sample. The authors took a percentile approach to the data instead of the traditional use of cutoff scores developed on the standardization group, recommending their approach as more appropriate and useful for clinicians seeing as the score distributions for their sample were fairly negatively skewed for various tests (Strickland et al. 2005).

Short Blessed Test The Short Blessed Test (Katzman et al. 1983) is a dementia screen that consists of six items that assess orientation to time, learning, delayed recall and the ability to count aloud backwards, and name the months of the year in reverse order. Scores on this measure are highly correlated with MMSE performance. Similar to the MMSE, this measure has been documented to have cultural biases (Hargrave 2006; Fillenbaum et al. 1998). In fact, Welsh et al. (1995) found a 62% false positive rate for dementia among African American participants compared to a 22% false–positive rate for White participants (Hargrave 2006). Thus, the short blessed memory test has poor sensitivity and specificity when administered to African American older adults and in light of the fact that other measures have great sensitivity, we do not recommend this measure for use with African Americans.

Mini-Cog-Functional Activities Questionnaire (MC-FAQ) The MC-FAQ is a screen that is often administered in primary care settings and takes just a few minutes to administer. The patient is asked to repeat three unrelated words, draw a clock, and then recall the three unrelated words (Pfeffer et al. 1982). In addition functional abilities are also assessed e.g., ability to write a check, prepare meals, etc. Tappen, Rosselli, and Engstrom (2012) found that age, years of education and depressive symptoms were significant predictors of classifications of normal, mild cognitive impairment, or dementia for the MC-FAQ. The researchers did not find, however, ethnic group differences on the classifications in their matched sample of 225 participants (75 African American, 75 Hispanic, and 75 European American adults; 59–95) suggesting that this measure may be culturally neutral.

Community Screening Interview for Dementia (CSI-D) The CSI-D is a 32 item cognitive test administered to the patient and a 26-item informant interview, enquiring after the participant's daily functioning and general health (Hall et al. 1993). The CSI-D has been utilized with culturally disparate, nonliterate, and literate populations in comparative epidemiological studies of dementia (World Health Organization 1992; Wessels et al. 2011). It has demonstrated good inter-rater reliability (99.4%), as well as specificity levels of 0.82–0.97, depending on the site, and good validity in detecting dementia in various populations (Wessels et al. 2011; Hall et al. 1996, 2000). Specific to this chapter, Shen et al. (2006) conducted an epidemiological study using the CSI-D in a large sample of 1493 African Americans aged 65 or older from Indianapolis, USA and 2459 elderly 65 years or older residing in Ibadan, Nigeria. The researchers concluded that informant reports are valid in assessing cognitive functioning, supporting the CSI-D as a culturally adaptable screening measure. Thus the CSI-D can be effectively used with African Americans. In addition to the above, there is research on this measure that is specific to African natives. Despite that the focus of this chapter is not on the African native population, this information is provided below as it is probable that a clinician looking for measures on Africans may look in this chapter and find this information useful. Chen et al. (2010) translated the CSI-D and administered it to a sample of 100 Kikuyu Kenyans aged 65 years or older, concluding that the specificity was nearly 90% for the detection of dementia if the sensitivity was assigned to be a 100%. The researchers determined that the CSI-D can be employed to detect dementia among East Africans.

AD8 The AD8 (Galvin et al. 2005) is an 8-item dementia screen. Items on the AD8 ask an informant to indicate whether or not there has been a change in their loved one's judgment, interest in hobbies and activities, forgetfulness etc. over the last several years. Answer options are "yes, a change"; "no, no change"; and "NA/don't know." Malstrom et al. (2009) administered the AD8 to 147 dyads aged 64–70 from the African American Health (AHH) project, finding that the measure alone was effective at discriminating between a clinical dementia rating of 0 and 0.5. The researchers concluded that the measure has high sensitivity and specificity for discerning clinical dementia ratings in the community and thus is recommended for use with African Americans.

Clinical Dementia Rating Scale The Clinical Dementia Rating (CDR) is a semi-structured interview that is conducted with a collateral contact to assess the identified client's memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. The person is assigned a score (on a 5-point spectrum) ranging from "none" to "severe." Scores in each of these are combined to obtain a composite score that ranges from 0 through 3 (Morris 1993). The CDR has been amply used in research with African Americans (i.e., Malstrom et al. 2009; Wilkins et al. 2006). Because the CDR (and others discussed herein) assess for Activities of Daily Living (ADLs) a discussion on cultural considerations when assessing for ADLs (and possibly Instrumental ADLs [IADLs]) is necessary.

ADLs and IADLs There are a number of factors that are specific to African Americans that can impact how African Americans are assessed when ADLs and IADLs are examined. In fact some cultural factors can result in diminished ADLs and IADLs. First, in terms of medical conditions that can impact ADLs and IADLs African Americans have high rates of diabetes, high blood pressure, lung disease, and obesity, which can contribute to an inability to perform tasks at a normative functioning level (Thorpe et al. 2009; Nies et al. 1999). Therefore the assessing clinician should be mindful of the increased risk factors for African Americans in terms of developing difficulties with ADLs and IADLs. Moreover, a number of sociodemographic factors including unsafe neighborhoods, limited space in household, poverty etc. can also contribute to difficulties in carrying out ADL's and IADL's such as exercising, grocery shopping, or maintaining a healthy diet (Fitzpatrick & Van Tran 1997). Furthermore, familial demands may make it difficult for African Americans to engage in self-care. For example, in a study of diabetic African American caregivers, it took them over 18 months to engage in self-care activities such as eye exams due to high family demands (Carthron et al. 2010). Higher levels of mental health diagnoses can also contribute to an inability to perform basic duties such as bathing, driving a car, or managing finances on one's own (Fitzpatrick &

Table 15.1 Summary of measures reviewed

Type of assessment	Test name	Useful to assess for	Research and recommendations
Screening	*Mini-Mental State Examination (MMSE)*	Establishing baseline cognitive performance; measuring cognitive decline	Fairly extensive research shows the MMSE has low specificity with African American elders, producing high rates of false positives. Research suggests that adjusting for literacy levels may correct this low specificity; Hawkins et al. (2011) provide age and education stratified normative data for African American elders
	Montreal Cognitive Assessment (MoCA)	Cognitive impairment	Normative data for the measure was derived from a large ($N=2653$), ethnically diverse, cardiac-diseased, population-based sample. In the absence of research on ethnic group comparisons of MoCA scores, it is of note that African Americans were included in the standardization sample (African Americans constituted 52 % of the sample). Recent research suggests caution is needed when interpreting MoCA scores in groups with lower education as well as those with lower literacy levels
	Cognistat	Cognitive decline	Research suggests that among African American elders (age 60+) there were not strong education and age effects on Cognistat performance. Strickland et al. (2005) provided age and education stratified normative data from their sample given lower performance levels in their sample of African American elders compared to the standardization group
	Short Blessed Test	Cognitive impairment	Fairly extensive research has demonstrated that the short blessed memory test has poor sensitivity and specificity when administered to African American older adults. The measure is not recommended for use with African Americans
	Mini-Cog–Functional Activities Questionnaire (MC-FAQ	Cognitive impairment	While age, years of education and depressive symptoms were significant predictors of score classifications, researchers' failure to find ethnic group differences suggests that this measure may be culturally neutral
Interview	*Community Screening Interview for Dementia (CSI-D)*[a] *AD-8*[a] *Clinical Dementia Rating Scale*[a] *ADLs and IADLs*[a]	Dementia Dementia Dementia Level of functioning	Comparative epidemiological research suggests that informant reports are valid in assessing cognitive functioning of African American elders, supporting the CSI-D as a culturally adaptable screening measure for this population Researchers have concluded that the measure has high sensitivity and specificity for discerning clinical dementia ratings in the community and thus is recommended for use with African Americans The CDR has been amply used in research with African Americans Some cultural factors can result in diminished ADLs and IADLs for older African Americans not specific to cognitive decline suggesting caution should be exercised when evaluating these

[a] This test can involve an informant interview in addition to the patient interview

Van Tran 1997). Thus, when assessing ADLs and IADLs the clinician should distinguish between limited functioning that is specific to cognitive decline vs. limited functioning that a result of the factors discussed above.

Summary and Recommendations

In this chapter we reviewed the literature on cognitive decline and dementia in African American elders with a specific focus on the Mini-Mental Status Examination, Montreal Cognitive Assessment (MoCA), Cognistat, Short Blessed Test, Mini-Cog-Functional Activities Questionnaire (MC-FAQ), Community Screening Interview for Dementia (CSI-D), AD-8, Clinical Dementia Rating (CDR), and examination of ADLS and IADLs (See Table 15.1 for a summary of these measures and relevant recommendations). While the focus of our chapter has been on cognitive decline (and mostly age related cognitive decline) a number of the measures discussed herein can be used to screen for cognitive impairment that either is or is not specific to dementia. In terms of recommendations, our first recommendation is that cutoff adjustments (e.g., those that account for age, education and literacy effects as discussed) be used when appropriate. Second, we recommend that literacy be factored into the evaluation. There is ample documentation that literacy levels can impact performance on the measures discussed in this chapter. Third, when a cognitive screen results in suspected cognitive decline, we recommend that more finite assessment be conducted given the high false–positive rate that screens can yield for African American individuals. Fourth, in order to inform assessments of premorbid functioning as well as inform the overall assessment process (e.g., early identification of areas to be assessed based upon family or caregiver reports of suspected impairment), we recommend collateral contacts occurring even at the screening stage. The clinician should keep in mind sociodemographic factors that impact ADLs and IADLs and where functioning in these domains is limited, it may be necessary to rule out other viable hypotheses for limited IADL/ADL functioning. This is particularly important as when age related cognitive decline is suspected medication may be prescribed and where false positives exist there can be iatrogenic effects. Lastly, for screens where the false positive rate is high (e.g., the MMSE) we recommend that performance on these screens be used in terms of establishing baseline cognitive performance and then used to track cognitive decline over time.

References

Administration on Aging. (2010). A statistical profile of black older Americans aged 65+. http://www.aoa.gov/Aging_Statistics/minority_aging/Facts-on-Black-Elderly-plain_format.aspx.

Aiken-Morgan, A. T., Sims, R., & Whitfield, K. E. (2010). Cardiovascular health and education as sources of individual variability in cognitive aging among African Americans. *Journal of Aging and Health, 22,* 477–503.

Alzheimer's Association. (n.d.). African Americans and Alzheimer's disease: The silent epidemic. www.alz.org/national/…/report_africanamericanssilentepidemic.pdf.

American Psychological Association. (2012). Guidelines for the evaluation of dementia and age-related cognitive change. *American Psychologist, 67*(1), 1–9.

Carthron, D. L., Johnson, T. M., Hubbart, T. D., Strickland, C., & Nance, K. (2010). "Give me some sugar!" The diabetes self-management activities of African-American primary caregiving grandmothers. *Journal of Nursing Scholarship, 42,* 330–337.

Chen, C.-H., Mizuno, T., Elston, R., Kariuki, M. M., Hall, K., Unverzagt, F., Hendrie, H., Gatere, S., Kioy, P., Patel, N. B., Friedland, R. P., & Kalaria, R. N. (2010). A comparative study to screen dementia and APOE genotypes in an ageing East African population. *Neurobiology of Aging, 31,* 732–740.

Chin, A. L., Negash, S., & Hamilton, R. (2011). Diversity and disparity in dementia: The impact of ethnoracial differences in Alzheimer Disease. *Alzheimer Disease and Associated Disorders, 25*(3), 187–195.

Clark, P. C., Kutner, N. G., Goldstein, F. C., Peterson-Hazen, S., Garner, V., Zhang, R., & Bowles, T. (2005). Impediments to timely diagnosis of Alzheimer's Disease in African Americans. *Journal of American Geriatric Society, 53*(11), 2012–2017.

Crowe, M., Clay, O. J., Sawyer, P., Crowther, M. R., & Allman, R. M. (2008). Education and reading ability in relation to differences in cognitive screening between African American and Caucasian older adults. *International Journal of Geriatric Psychiatry, 23,* 222–223.

Dong, Y., Sharma, V. K., Chan, B. P., Venketasubramanian, N., Teoh, H. L., Seet, R. C., Tanicala, S., Chan, Y. H., & Chen, C. (2010). The Montreal Cognitive Assessment (MoCA) is superior to the Mini-Mental

State Examination (MMSE) for the detection of vascular cognitive impairment after acute stroke. *Journal of Neuroscience, 299*(1), 15–18.

Ferman, T. J., Lucas, J. A., Ivnik, R. J., Smith, G. E., Willis, F. B., Petersen, R. C., et al. (2005). Mayo's older African American normative studies. Auditory Verbal Learning Test norms for African American elders.

Fillenbaum, G. G., Landerman, L., & Simonsick, E. M. (1998). Equivalence of two screens of cognitive functioning: The short portable mental status exam and the orientation-memory-concentration test. *Journal of American Geriatric Society, 46,* 1512–1518.

Fitzpatrick, T. R., & Van Tran, T. (1997). Age, gender, and health among African Americans. *Social Work in Health Care, 26*, 69–85.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198.

Froehlich, T. E., Bogardus, S. T., & Inouye, S. K. (2001). Dementia and race: Are there differences between African Americans and Caucasians? *Journal of American Geriatrics Society, 49,* 477–484.

Galvin, J. E., Roe, C. M., Powlishta, K. K., Coats, M. A., Muich, S. J., Grant, E., Miller, M., Storandt, M., & Morris, J. C. (2005). The AD8: A brief informant interview to detect dementia. *Neurology, 65*(4), 559–564.

Griffith, P., Lichtenberg, P., Goldman, R., & Payne-Parrish, J. (2006). Safety and efficacy of Donepezil in African Americans with mild-to-moderate Alzheimer's Disease. *Journal of the National Medical Association, 98*(10), 1590–1597.

Hall, K. S., Hendrie, H. H., Brittain, H. M., Norton, J. A., Rodgers, D. D., Prince, C. S., Pillay, N., Blue, A. W., Kaufert, J. N., Nath, A., Shelton, P., Postrl, B. D., & Osuntokun, B. O. (1993). The development of a dementia screening interview in two distinct languages. *International Journal of Methods in Psychiatric Research, 3,* 1–28.

Hall, K. S., Ogunniyi, A. O., Hendrie, H. C., Osuntokun, B. O., Hui, S. L., Musick, B. S., Rodenberg, C. A., Unverzagt, F. W., Guerje, O., & Bajyewu, O. (1996). Study of dementias: Methods and performance of the survey instrument Indianapolis, U.S.A., and Ibadan, Nigeria. *International Journal of Methods in Psychiatric Research, 6,* 129–142.

Hall, K. S., Gao, S., Emsley, C. L., Ogunniyi, A. O., Morgan, O., & Hendrie, H. C. (2000). Community screening interview for dementia (CSI 'D'); performance in five disparate study sites. *International Journal of Geriatric Psychiatry, 15,* 521–531.

Hargrave, R. (2006). Neurocognitive assessment of dementia in African American elders. In G. Yeo & D. Gallagher-Thompson (Eds.), *Ethnicity and the Dementias* (2nd ed, pp. 33–53). Boston: Routledge.

Hargrave, R. (2010). Health and Health Care of African American Older Adults. http://geriatrics.stanford.edu/ethnomed/african_american/downloads/african_american.pdf.

Hawkins, K. A., Cromer, J. R., Piotrowski, A. S., & Pearlson, G. D. (2011). Mini-Mental State Exam performance of older African Americans: Effect of age, gender, education, hypertension, diabetes, and the inclusion of serial 7 s subtraction versus "world" backward on score. *Archives of Clinical Neuropsychology, 26,* 645–652.

Innes, A. (2009). *Dementia studies within cultural contexts. In Dementia studies: A social science perspective* (pp. 72–101). Thousand Oaks: Sage.

Jett, K. F. (2006). Mind-loss in the African American community: Dementia as a normal part of aging. *Journal of Aging Studies, 20*(1), 1–10.

Julayanont, P., Hemrungrojn, S., & Tangwongchai. S. (2013). The effect of education and literacy on performance on the Montreal Cognitive Assessment among cognitively normal elderly. Alzheimer's and Dementia: *The Journal of the Alzheimer's Association, 9*(4), 793.

Katzman, R., Brown, T., Fuld, P., Peck, A., Schechter, R., & Schimmel, H. (1983). Validation of a short orientation-memory-concentration test of cognitive impairment. *American Journal of Psychiatry, 140*(6), 734–739.

Kiernan, R. J., Mueller, J., Langston, J. W., & Van Dyke, C. (1987). The Neurobehavioral Cognitive Status Examination: A brief but differentiated approach to cognitive assessment. *Annals of Internal Medicine, 107,* 481–485.

Lucas, J. A., Ivnik, R. J., Smith, G. E., Ferman, T. J., Willis, F. B., Petersen, R. C., et al. (2005a). A brief report on WAIS-R normative data collection in Mayo's Older African Americans Normative Studies. *The Clinical Neuropsychologist, 19*(2), 184–188.

Lucas, J. A., Ivnik, R. J., Smith, G. E., Ferman, T. J., Willis, F. B., Petersen, R. C., et al. (2005b). Mayo's Older African American Normative Studies: Norms for Boston Naming Test, Controlled Oral Word Association, Category Fluency, Animal Naming, Token Test, WRAT-3 Reading, Trail Making Test, Stroop Test, and Judgment of Line Orientation. *The Clinical Neuropsychologist, 19*(2), 43–269.

Lucas, J. A., Ivnik, R. J., Willis, F. B., Ferman, T. J., Smith, G. E., Parfitt, F. C., et al. (2005c). Mayo's Older African Americans Normative Studies: WMS-R norms for African American elders. *The Clinical Neuropsychologist, 19*(2), 189–213.

Lucas, J. A., Ivnik, R. J., Willis, F. B., Ferman, T. J., Smith, G. E., Parfitt, F. C., et al. (2005d). Mayo's Older African Americans Normative Studies: Normative data for commonly used clinical neuropsychological measures. *The Clinical Neuropsychologist, 19*(2), 162–183.

Malstrom, T. K., Miller, D. K., Coats, M. A., Jackson, P., Miller, J. P., & Morris, J. C. (2009). Informant-based dementia screening in a population-based sample of African Americans. *Alzheimer Disease and Associated Disorders, 23*(2), 117–123.

Morgan, A. A., Marsiske, M., & Whitfield, K. E. (2008). Characterizing and explaining differences in cognitive test performance between African American and

European American older adults. *Experimental Aging Research, 34,* 80–100.

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology, 43*(11), 2412–2414.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment (MoCA): A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53,* 695–699.

Nies, M. A., Vollman, M., & Cook, T. (1999). African American women's experiences with physical activity in their daily lives. *Public Health Nursing, 16,* 23–36. doi:10.1046/j.1525-1446.1999.00023.x.

Osato, S. S., Yang, J., & La Rue, A. (1993). The Neurobehavioral Cognitive Status Examination in an older psychiatric population: An exploratory study of validity. *Neuropsychiatry Neuropsychology, and Behavioral Neurology, 6,* 103–110.

Osato, S. S., La Rue, A., & Yang, J. (1994). Validity of the Neurobehavioral Cognitive Status Examination (NCSE) subtests in older psychiatric patients. *Journal of the International Neuropsychological Society, 1,* 356.

Pedraza, O., Clark, J. H., O'Bryant, S. E., Smith, G. E., Ivnik, R. J., Graff-Radford, N. R., Willis, F. B., Petersen, R. C., & Lucas, J. A. (2012). Diagnostic validity of age and education corrections for the Mini-Mental State Examination in older African Americans. *Journal of the American Geriatrics Society, 60*(2), 328–331.

Pfeffer, R. I., Kurosaki, T. T., Harrah, Jr. C. H., Chance, J. M., & Filos. S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology, 37*(3), 323–329.

Rilling, L., M., Lucas, J. A., Ivnik, R. J., Smith, G. E., Willis, F. B., Ferman, T. J., et al. (2005). Mayo's Older African American Normative Studies: Norms for the Mattis Dementia Rating Scale. *Clinical Neuropsychology, 19,* 229–242.

Rossetti, H. C., Lacritz, L. H., Cullum, C. M., & Weiner, M. F. (2011). Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology, 77,* 1272–1275.

Schneider, B. C., & Lichtenberg, P. A. (2011). Influence of reading ability on neuropsychological performance in African American elders. *Archives of Clinical Neuropsychology, 26,* 624–631.

Shen, J., Gao, S., Unverzagt, F. W., Ogunniyi, A., Baiyewu, O., Gureje, O., Hendrie, H. C., & Hall, K. S. (2006). Validation analysis of informant's ratings of cognitive function in African Americans and Nigerians. *International Journal of Geriatric Psychiatry, 21,* 618–625.

Sims, R., Whitfield, K., Gamaldo, A., Edwards, C., & Allaire, J. (2008). Self-perception of memory versus memory performance in older African Americans: The role of well-being. *The Gerontologist, 48*(S3), 408.

Strickland, T. L., Longobardi, P. G., Alperson, B. L., & Angre, K. (2005). Mini-Mental State and Cognistat performance in an older African American sample. *The Clinical Neuropsychologist, 19,* 87–98.

Tappen, R. M., Rosselli, M., & Engstrom, G. (2012). Use of the MC-FAQ and MMSE-FAQ in cognitive screening of older African Americans, Hispanic Americans, and European Americans. *American Journal of Geriatric Psychiatry, 20*(11), 955–962.

Thorpe, R. J., Szanton S. L., & Whitfield, K. (2009). Association between lung function and disability in African-Americans. *Journal of Epidemiology and Community Health, 63*(7), 541–545.

Welsh, F. G., Wilkinson, W., Heyman, A., Mohs, R. C., Stern, Y., Harrell, L., et al. (1995). Neuropsychological test performance in African American and white patients with Alzheimer's disease. *Neurology, 45,* 207–2211.

Wessels, A. M., Lane, K. A., Gao, S., Hall, K. S., Unverzagt, F. W., & Hendrie, H. C. (2011). Diabetes and cognitive decline in elderly African Americans: A 15-year follow-up study. *Alzheimer's and Dementia, 7,* 418–424.

Wilkins, C. H., et al. (2006). The neuropathology of Alzheimer disease in African American and white individuals. *Archives of Neurology, 63*(1), 87–90.

World Health Organization (1992). *International classification of disease, 10th Revision (ICD-10)*. Geneva: WHO.

Assessing Psychosis in African American Clients

16

Gregory P. Strauss and Adam J. Culbreth

Overview

Schizophrenia is a heterogeneous disorder that has been documented in nearly every culture in the world (Jablensky et al. 1992). However, there are reasons to believe that the symptom patterns associated with the syndrome of schizophrenia do not manifest identically across cultures or people of different ethnicities (Myers 2011). Of particular interest, there is a large body of evidence suggesting that the prevalence of schizophrenia is greater in African Americans than Caucasians (Bresnahan et al. 2007; Lipton and Simon 1985; Liss et al. 1973; Neighbors et al. 1999; Strakowski et al. 1996a), with some evidence suggesting that African Americans may be up to three times more likely to be diagnosed with schizophrenia than Caucasians (Bresnahan et al. 2007). Initially, researchers believed that factors other than ethnicity, such as lower socioeconomic status or clinicians not strictly adhering to structured clinical diagnostic interviews, may account for the increased prevalence among African Americans. However, after taking socioeconomic status into consideration and attempting to strictly adhere to diagnostic manuals, studies have still found greater prevalence of schizophrenia diagnoses among African American populations, although this difference is notably attenuated (Bresnahan et al. 2007). The reasons for this increase are currently unknown. However, the absence of simple explanations has caused researchers to focus on multifaceted accounts that take a myriad of contextual risk factors into consideration, such as: immigration, cumulative social disadvantage, adverse life events, and ethnic density. Diagnostic biases and lack of cultural consideration have also been posed as viable explanations for the increased rate of schizophrenia diagnoses among African Americans, with several studies suggesting that clinicians may misunderstand the cultural salience of psychotic symptoms in African Americans with psychosis (Adebimpe 1981; Adebimpe et al. 1982). If true, the misdiagnosis of schizophrenia in the African American population would prove to be a substantial problem, as effective treatments for schizophrenia are rarely similar to other psychiatric illnesses and the stigma of being misdiagnosed with schizophrenia may negatively affect these individuals and their families. These issues have sparked a debate within the scientific literature regarding the validity of diagnostic and assessment procedures in African Americans. In the current chapter, we review this literature on diagnosing psychotic disorders in African American clients, highlighting

G. P. Strauss (✉)
Psychology Department, State University of New York at Binghamton, PO Box 6000, Binghamton, NY 13902-6000, USA
e-mail: gstrauss@binghamton.edu

A. J. Culbreth
Department of Psychology, Washington University in St. Louis, Box 1125, One Brookings Drive, St. Louis, MO 63130, USA
e-mail: aculbreth@wustl.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_16,

the leading theories regarding the increased prevalence of psychotic disorder diagnoses in African Americans. Given the importance of symptom severity assessment in diagnosis, we also provide a summary of major clinical rating scales and review the literature on the assessment of positive, negative, and disorganized symptoms in African Americans. We also present new data on two of the most common clinical rating scales used to assess these symptoms since little published data exists regarding psychometric properties of major rating scales in African American and Caucasian clients separately, and conclude with recommendations regarding the assessment of psychosis in African American clients.

Diagnosing Psychotic Disorders in African American Clients

Diagnostic Errors and Rater Bias A substantial number of studies have demonstrated that schizophrenia is more often diagnosed among African American than Caucasian clients, while mood disorders are more frequently diagnosed in Caucasians than African Americans (Lawson 1986; Neighbors et al. 1989; Simon et al. 1973; Strakowski et al. 1993, 1996b; Worthington 1992). Such findings have led some to question the accuracy of clinical judgment, and propose that the elevated rates of schizophrenia in African Americans reflect diagnostic errors and rater biases that result from failure to consider cultural factors during diagnostic and assessment procedures (Neighbors et al. 2003). Consistent with this notion, studies have found that clinicians use different criteria to diagnose schizophrenia in Caucasian Americans and African Americans (Trierweiler et al. 2006). For instance, African Americans are more likely, than Caucasian Americans, to receive a diagnosis of schizophrenia when evaluated in a hospital setting; however, when clinicians reevaluate those same patients using semi-structured diagnostic rating instruments that are linked to diagnostic and statistical manual (DSM) criteria, diagnostic race differences are markedly reduced (Neighbors et al. 1999, 2003). This may imply that clinical diagnoses made without the use of structured diagnostic procedures are less accurate than those that do use structured diagnostic tools like the Structured Clinical Interview for DSM-IV (SCID; First et al. 2002). Alternatively, clinicians may not be sensitive to racial and ethnic differences in symptom expression in African Americans. For example, Neighbors et al. (2003) found evidence that clinicians use different processes to link symptom observations to diagnostic categories in individuals of different ethnicities. In particular, loose associations, inappropriate affect, auditory hallucinations, and vague speech predicted receipt of schizophrenia diagnosis in African Americans to a greater extent than Caucasians. Although the frequency of these symptoms did not differ between ethnic groups, the weight that clinicians ascribed to them did differ, and subsequently influenced whether a mood or psychotic disorder diagnosis was assigned. Furthermore, the culture of the diagnosing clinician has been found to influence the extent to which different symptom clusters are emphasized when making the schizophrenia diagnosis. African American clinicians are more likely to emphasize positive symptoms than non-African American clinicians, and only non-African American clinicians tend to emphasize negative symptoms (e.g., poverty of speech, blunted facial affect) in the schizophrenia diagnosis (Trierweiler et al. 2006). This suggests that when evaluating African Americans, non-African American clinicians may be more influenced by barriers to communication than their African American colleagues. In addition to causing blurred boundaries between different aspects of schizophrenia psychopathology, it is possible that lack of cultural consideration also results in higher rates of misclassification of schizophrenia in African Americans and for the disorder as a whole.

A number of studies have also reported that African Americans diagnosed with schizophrenia report experiencing more severe psychotic symptoms (Adebimpe et al. 1982; Mukherjee et al. 1983) and a greater number of “first-rank” psychotic symptoms than Caucasian patients (Strakowski et al. 1996a, b; Arnold et al. 2004). In part, this difference in symptom presentation may be due to the fact that African Americans tend to make greater use of emergency centers where clinicians are likely to see symptoms at their highest severity, thereby making them more prone to assigning a diagnosis of schizophrenia when such symptoms are reported (Trierweiler et al. 2006). In addition, although Schneiderian first-rank symptoms have historically been considered by many clinicians to be the hallmark symptomatology of schizophrenia, and many clinicians believe that these symptoms are the most valuable markers of pathology when making a schizophrenia diagnosis (Schneider 1959), research suggests that the presence of first-rank symptoms may not be as predictive of schizophrenia as had been believed in years past (Carpenter et al. 1973; Andreasen & Flaum 1991). Such evidence has lead to changes in the diagnostic criteria for schizophrenia in the DSM-5, and first-rank symptoms will no longer be weighted more heavily in allowing subjects to meet DSM criteria in the absence of other required symptoms of schizophrenia (Keller et al. 2011). It is therefore possible that greater reliance on first-rank symptoms, when diagnosing African Americans, leads to higher prevalence rates. It will be important to determine whether these DSM-5 changes alter the increased prevalence of schizophrenia diagnoses in African Americans relative to Caucasians.

Overall, the aforementioned findings may suggest that the higher prevalence of psychotic disorder diagnoses in African Americans results at least to some extent due to a lack of cultural consideration during diagnostic interviewing. However, there is currently no definitive evidence that the elevated incidence of schizophrenia in African Americans is solely due to widespread biases or errors in clinical diagnoses. Indeed, recent meta-analyses indicating increased incidence of schizophrenia diagnosis across a range of ethnic groups, particularly migrant groups, suggest that sociocultural and other factors may also be at play.

Immigration Many studies have suggested that the increased prevalence of schizophrenia among minorities is due to the myriad of environmental and sociocultural factors that accompany immigration (Odegaard 1932; Bourque et al. 2011; Gara et al. 2012; Cantor-Graae and Selten 2005; Cantor-Graae and Pedersen 2007). It is well-documented that first-generation migrants are at an increased risk for developing schizophrenia, and this discrepancy is still observable in second-generation migrants (Bourque et al. 2011). However, these differences in prevalence are not witnessed as robustly or reliably among all immigrants or minority populations. For example, Latin Americans do not differ from Caucasians in diagnostic prevalence of schizophrenia (Gara et al. 2012; Minsky et al. 2003), suggesting that sociocultural factors beyond immigration may also be at play. Of particular relevance to African American populations, individuals who migrated from a country where dark skin color is present in the majority to a country where white skin color is the majority show a greater prevalence of schizophrenia diagnoses than other migrant groups (Cantor-Graae and Selten 2005; Gara et al. 2012). Some have attributed this discrepancy to the fact that dark-skinned individuals are more readily discriminated as compared to other migrant groups due to their clearly observable difference from the native people (Murray and Hutchinson 1999; Sharply et al. 2001; Cantor-Graae and Selten 2005). Along with this thinking, Selten and Cantor-Graae suggest that the constant and prolonged experience of having an outsider status in a new country may cause dark-skinned individuals to possess a chronic stressor that could be a contributing factor to the higher rates of schizophrenia amongst dark-skinned migrant populations (2005; 2007).

Urban Density, Socioeconomic Status, and Social Disadvantage Several additional sociocultural factors may influence the increased prevalence of schizophrenia in African Americans. Ethnic density has been found to be associated with higher rates of schizophrenia in multiple cultures. For example, in a study conducted in the United Kingdom, it was found that when individuals with black skin made up less than 25 % of the population of their neighborhood there was approximately three times greater risk for developing schizophrenia. However, this risk became nonsignificant if the neighborhood consisted of 25% or more individuals with black skin (Schofield et al. 2010). This suggests that risk for psychotic disorders is associated with one's level of acculturation, as well as how isolated they are from the culture that they identify with. Urbanicity and socioeconomic status have also been linked to higher rates of schizophrenia diagnoses, especially lower socioeconomic status during childhood (Corcoran et al. 2009; Jenkins et al. 2008; March et al. 2008). Increased risk for traumatic and adverse life events has been one potential link between psychosis and social disadvantage, potentially compounding likelihood of developing psychosis if individuals are at genetic risk (Arsneault et al. 2010). Collectively, these findings indicate that a range of sociocultural factors may contribute to the increased prevalence of schizophrenia in African American clients.

Assessment of Positive, Negative, and Disorganized Symptoms in Individuals with Psychotic Disorders

Although there is considerable debate regarding the reasons underlying the increased prevalence of schizophrenia diagnoses in African Americans, relatively little research has examined differences in the psychometric properties of major scales used to assess the symptoms associated with schizophrenia. It would be important to know how these scales function in African Americans since these measures are commonly used to test the efficacy of new medications in clinical trials and to monitor changes in symptom severity in clinical practice. In the remainder of this chapter, we provide a summary of the clinical rating instruments that are most commonly used to assess symptoms of schizophrenia and other psychotic disorders in research and clinical settings. Given the paucity of available data on African Americans specifically, we also present previously unpublished archival data on the psychometric properties of several of the most common schizophrenia symptom rating instruments in samples of Caucasian and African American individuals diagnosed with a psychotic disorder.

As previously noted, schizophrenia is a markedly heterogeneous disorder with regard to its symptom presentation. Modern factor analytic studies typically support the existence of three major domains of psychopathology in schizophrenia: positive, negative, and disorganized symptoms (Keefe et al. 1992; Kelley et al. 1999; Mueser et al. 1994; Peralta and Cuesta 1995; Sayers et al. 1996). Within these broad symptom cluster distinctions, there is also evidence for separate individual symptom dimensions. For example, positive symptoms are typically divided into hallucinations and delusions, and disorganization into formal thought disorder and bizarre behavior (Andreasen et al. 1995). Negative symptoms are also multidimensional with consistent evidence for two dimensions reflecting motivation and pleasure (e.g., anhedonia, avolition, asociality) and emotional expressivity (e.g., alogia, restricted affect) (Blanchard and Cohen 2006; Kirkpatrick et al. 2011; Strauss et al. 2012, 2013).

A variety of measures have been developed to assess positive, negative, and disorganized symptoms in individuals with psychotic disorders. Table 16.1 presents the clinical assessment tools most commonly used to measure positive, negative, and disorganized symptoms in the field. A description of each measure is included within the table. These symptom severity measures are typically completed by clinicians who perform a standard clinical interview designed to assess the relevant symptom domains, and then rate each item on the scale on the basis of their observations, patient self-report, and sometimes collateral report. The scoring procedures for these measures vary; however, most of these scales are scored by totaling all items on the scale or by totaling the individual items that form its subscales. Unlike many psychological tests, these psychiatric clinical rating scales typically do not have established norms for estimating standard scores or severity percentiles. Although norms do not exist, these scales are typically thought to be valid for use in individuals of different ages, stages of illness, and cultural groups.

Table 16.1 Summary of major symptom instruments used to rate positive, negative, and disorganized symptoms in psychotic disorders

Name	Author(s)/Date	Symptom domains assessed	Description
Positive and Negative Syndrome Scale (PANSS)	Kay et al. 1987	General psychiatric; positive, negative	The PANSS is a 30-item scale used to assess schizophrenia across three domains: positive symptoms, negative symptoms, and general severity of mental illness. Individual items are rated on a 7-point Likert type scale. Global ratings are used to represent the overall severity of the symptoms within each of the three domains
Brief Psychiatric Rating Scale (BPRS)	Overall and Gorham 1962	General psychiatric; positive, negative; disorganized	The original BPRS scale consisted of 16 items (Overall and Gorham 1962), which are rated on a 7-point Likert scale. More recent versions consist of 18 (Overall and Gorham 1988) or 20 items (Lukoff et al. 1986). Individual items assess positive, negative, disorganized, and general psychiatric symptoms
Scale for the Assessment of Positive Symptoms (SAPS)	Andreasen 1984	Positive; disorganized	The SAPS is a 35-item scale measuring positive and disorganized symptoms in four primary domains: hallucinations, delusions, bizarre behavior, and positive formal thought disorder. Symptoms are typically rated over the past week on a 6-point scale. Global ratings are used to represent overall severity within each of these five domains, taking into account both the nature and severity of all symptoms observed
Psychotic Symptom Rating Scales (PSYRATS)	Haddock et al. 1999	Positive	The PSYRATS is a 17-item scale measuring the presence and severity of auditory psychosis. The scale is divided into two subscales: hallucinations and delusions. Individual items on each subscale are rated on a 5-point ordinal scale. Global ratings are used to the severity of symptoms within a given subscale. The PSYRATS has the advantage of being able to assess multiple dimensions of auditory psychosis
Scale for the Assessment of Negative Symptoms (SANS)	Andreasen 1983	Negative	The original SANS consisted of 30 items designed to assess negative symptom domains such as blunted affect, alogia, avolition, anhedonia, asociality, and attention. More recent versions have reduced the number of items to 25 or 22, excluding items related to attention, poverty of content of speech, etc. Symptoms are rated on a 6-point scale, and typically evaluated over a 1-week or 1-month period. A global rating is also made for each core domain that takes into account the nature and severity of items within that scale
Negative Symptom Assessment (NSA)	Axelrod et al. 1993	Negative	The original NSA is a 16-item scale used to assess the negative symptoms of schizophrenia. Individual items on the NSA are rated using a 7-point Likert-type scale. Recently the number of items of the NSA has been reduced from 16 to 4 (Alphs et al. 2011). A global rating is calculated to assess the individual's degree of negative symptom severity compared to a healthy individual

Table 16.1 (continued)

Name	Author(s)/Date	Symptom domains assessed	Description
Brief negative symptom scale (BNSS)	Kirkpatrick et al. 2011	Negative	The BNSS is a 13-item scale designed to assess the severity of anhedonia, asociality, avolition, alogia, restricted affect, and lack of normal distress. Item severity is rated on a 7-point scale over the past week timeframe. The BNSS has advantages over existing measures in that it evaluates multiple components of pleasure (e.g., frequency, retrospective, prospective), as well as dissociations between internal experience and outward behavior for avolition and asociality
Clinical Assessment Interview for Negative Symptoms (CAINS)	Kring et al. 2013	Negative	The CAINS is a 13-item scale designed to assess the domains of anhedonia, asociality, avolition, alogia, and restricted affect. It offers the advantage of assessing the frequency of past week pleasure, and the anticipated frequency of future pleasure
Schedule for the Deficit Syndrome (SDS)	Kirkpatrick et al. 1989	Negative	The Schedule for the Deficit Syndrome (SDS) is used to classify patients according to deficit/non-deficit status (i.e., whether they have primary and enduring negative symptoms or not). The SDS requires a semi-structured clinical interview designed to assess severity of negative symptoms in relation to six domains: restricted affect, diminished emotional range, poverty of speech, curbed interests, diminished sense of purpose, and diminished social drive. Severity ratings are made on a 5-point rating scale. For each symptom domain, symptoms are further classified as being primary/secondary (i.e., idiopathic, not due to secondary negative symptom factors) and stable/unstable (lasting >1 year). To be classified as a deficit syndrome case, patients must: (1) meet DSM criteria for schizophrenia, (2) evidence moderate or higher (SDS severity of 2 or >) symptom severity on at least two of the six symptom domains, (3) have at least two of these symptoms considered primary, and (4) demonstrate a stable symptom presentation during periods of relative remission over the past year

Information regarding cultural considerations in the use of these measures with African American clients is lacking. Our data reported here on the SAPS and SANS suggest that these scales have good reliability and validity for use in African Americans with schizophrenia, and that they may not require adaptation. However, as a general rule-of-thumb, positive, negative, and disorganized symptoms of schizophrenia should be evaluated in relation to cultural context when performing a clinical or diagnostic interview

It has yet to be empirically determined whether the assumption that these scales have cross-cultural utility is correct. Very few published studies have evaluated ethnicity-related differences in major clinical measures across cultures, including African American clients. Of the few studies that have been conducted, results indicate that African Americans are rated as having more first-rank symptoms on the SAPS (Arnold et al. 2004), and more severe positive symptoms on select items of the PANSS (suspiciousness and hallucinatory behavior) (Barrio et al. 2003); however, there are typically no overall differences on broad positive, negative, or disorganized symptom domain scores on the PANSS, or total negative symptom scores on the SANS (Arnold et al. 2004; Barrio et al. 2003) between African Americans and Caucasians. It is currently unclear whether the psychometric properties of these instruments differ between African Americans and Caucasians; however, it would be important to examine differences in reliability and validity given the aforementioned evidence related to increased prevalence of schizophrenia diagnoses in African Americans and potential issues surrounding rater bias.

Evaluation of Psychometric Properties of Major Clinical Rating Scales in African American and Caucasian Clients

Within our group at the Maryland Psychiatric Research Center (MPRC) at the University of Maryland School of Medicine, we are in a unique position to evaluate the psychometric properties of major assessments used to index positive, negative, and disorganized symptoms in African American clients diagnosed with psychotic disorders. Our outpatient and inpatient research units at the MPRC focus exclusively on the etiology and treatment of schizophrenia. We have collected index admission data on hundreds of individuals over the past 25 years, including data on major psychiatric rating scales. In the sections that follow, we present data on the reliability and validity of the two most popular instruments used in the assessment of psychosis, with reliability and validity analyses conducted separately for African American and Caucasian subjects meeting criteria for a DSM diagnosis of a psychotic disorder. Measures evaluated include the SAPS (Andreasen 1984) and the SANS (Andreasen 1983). The content of these measures is outlined in Table 16.1. All participants tested in our index assessments provided written informed consent for a protocol approved by the University of Maryland.

For each measure, analyses focused on: (1) *Reliability*: evaluated in relation to internal consistency and alpha-if-item-deleted analyses; (2) *Construct validity*: evaluated via principal components analysis (varimax rotation with Kaiser normalization) to evaluate the internal structure of the scales; (3) *Convergent validity*: evaluated via bivariate correlations with measures purported to index similar symptom domains on the Brief Psychiatric Rating Scale (BPRS: Overall and Gorham 1962); (4) *Discriminant validity*: evaluated via bivariate correlations with measures thought to index symptom constructs with minimal to moderate relationships with the measure of interest on the BPRS (Overall and Gorham 1962), and (5) *Basic descriptive statistics:* differences in symptom severity between Caucasian and African American subjects were compared using ANOVA.

1. Scale for the Assessment of Positive Symptoms (SAPS) 1A. Reliability: Cronbach's alpha, calculated to examine internal consistency, was excellent for African American and Caucasian patients (see Table 16.2), indicating that the SAPS items measure a single latent construct of positive symptoms in both ethnic groups. In addition, alpha if-item-deleted coefficients were high in both ethnic groups, suggesting no benefit from excluding any individual items from the total score (see Table 16.2). Thus, the SAPS demonstrated good reliability in African American and Caucasian subjects meeting diagnostic criteria for psychotic disorders.

1B. Construct Validity: Principal components analysis was used to examine the factor structure of the SAPS. Results indicated a 2-factor solution for Caucasians, and a 1-factor solution for African Americans (see Table 16.3). The 2-factor solution seen in Caucasians is consistent with prior factor analytic work on the SAPS (Andreasen et al. 1995). These factors reflect psychosis and disorganization symptom dimensions. Evidence for a single factor in African American subjects may reflect rater bias, whereby clinicians tend to rate both psychotic and disorganized symptom dimensions similarly in African Americans, but perceive differences in Caucasians. Alternatively, the single factor may reflect genuine differences in symptom expression, such that psychosis and disorganization tend to travel together in African Americans more frequently than Caucasians.

Table 16.2 Scale psychometrics: SAPS—reliability analyses

	Caucasian (n=239)	African American (n=180)
Mean SAPS global scores (SD)		
Hallucinations	2.13 (1.93)	2.18 (2.03)
Delusions	2.43 (1.71)	2.51 (1.66)
Bizarre behavior	1.03 (1.30)	0.89 (1.21)
Thought disorder	1.30 (1.46)	1.27 (1.30)
Cronbach's alpha	0.90	0.92
Alpha-if-item deleted range	0.89–0.90	0.91–0.92

Table 16.3 Scale psychometrics: SAPS—factor analyses

	Caucasian (n=239)		African American (n=180)
Global item	Factor 1	Factor 2	Factor 1
Hallucinations	*0.89*	0.08	*0.75*
Delusions	*0.87*	0.18	*0.81*
Bizarre behavior	0.18	*0.82*	*0.63*
Thought disorder	0.07	*0.85*	*0.70*
Eigen value	1.96	1.07	2.11
% Variance	49.0	26.7	52.8

Table 16.4 Scale psychometrics: SAPS—convergent and discriminant validity

	Caucasian	African American
BPRS positive symptoms	0.68***	0.59***
BPRS negative symptoms	0.09	−0.09
BPRS disorganized symptoms	0.28***	0.49***
BPRS total symptoms	0.56***	0.45***

* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

1C. Convergent Validity: In Caucasians and African Americans the SAPS total score was highly correlated with the BPRS psychosis score, suggesting that the SAPS has good convergent validity with another established measure of positive symptoms (see Table 16.4). Although the correlation between SAPS and BPRS positive scores was higher for Caucasians than African Americans, the test for significant differences between correlations indicated that this difference was nonsignificant. However, the test for significant differences in correlations was significant between Caucasians and African Americans with regard to the relationship between SAPS total and BPRS disorganization. This suggests that the SAPS may have better convergent validity in African Americans than Caucasians. This conclusion is strengthened by the fact that the BPRS disorganized dimension was more highly correlated with the SAPS global formal thought disorder item on the SAPS in African Americans (r=0.43) than Caucasians (r=0.30).

1D. Discriminant Validity: A comparison of correlations between the SAPS total score and the BPRS Positive, Disorganized, Negative, and Total symptom subscale scores supported the discriminant validity of the SAPS in Caucasians and African Americans (see Table 16.4).

1E. Comparison of Mean SAPS Scores: One-way ANOVAs calculated separately for the four SAPS global scores indicated that Caucasians and African Americans did not significantly differ in positive or disorganized symptom severity (all p's > 0.27) (see Table 16.2).

2. Scale for the Assessment of Negative Symptoms 2A. Reliability: On the SANS, Cronbach's alpha was good for African Americans, Caucasians, and all subjects (see Table 16.5); however, alpha was slightly lower for African

Table 16.5 Scale psychometrics: SANS—reliability analyses

	Caucasian (n = 155)	African American (n = 136)
Mean SAPS global scores (SD)		
Affective blunting	1.76 (1.16)	1.65 (1.24)
Alogia	1.05 (1.08)	1.05 (1.06)
Avolition	2.50 (1.29)	2.40 (1.36)
Anhedonia-Asociality	2.32 (0.94)	2.26 (1.13)
Cronbach's alpha	0.89	0.84
Alpha-if-item deleted range	0.88–0.90	0.83–0.85

Table 16.6 Scale psychometrics: SANS—factor analyses

	Caucasian (n = 155)	African American (n = 136)	
Global item	Factor 1	Factor 1	Factor 2
Affective Blunting	*0.77*	*0.78*	0.21
Alogia	*0.62*	*0.60*	0.14
Avolition	*0.64*	0.11	*0.63*
Anhedonia-Asociality	*0.63*	0.26	*0.72*
Eigen value	2.32	2.00	1.02
% Variance	58.0	49.9	25.5

Americans than Caucasians. In addition, alpha if-item-deleted coefficients were good in both ethnic groups, suggesting no evidence for poor reliability among any individual SANS items (see Table 16.5), although these were slightly lower in African Americans. Overall, these findings suggest good internal consistency among SANS items in Caucasians, and slightly lower but still good internal consistency in African Americans.

2B. Construct Validity: Principal components analysis with maximum-likelihood rotation was used to examine the factor structure of the SANS. Results indicated a 1-factor solution in all subjects and Caucasians, and a 2-factor solution in African Americans (see Table 16.6). The 2-factor solution found in African Americans is the one most commonly found on the SANS, as well as other negative symptom measures (Blanchard and Cohen 2006; Strauss et al. 2012, 2013), with factors representing motivation and pleasure (avolition, anhedonia, asociality) and diminished expression (affective blunting, alogia). It is possible that prior factor analytic results on the SANS and other measures have primarily been driven by African American subjects. The differential factor structure of the SANS in African American and Caucasian subjects may reflect a valid difference in symptom presentation, such that diminished expression and motivation/pleasure tend to travel together in Caucasians, but not in African Americans. In a recent study by Strauss et al. (2013), it was found that schizophrenia patients could be separated into distinct negative symptom subgroups based upon the relative severity of their diminished expression and motivation/pleasure scores. Separable groups of patients with relatively higher scores on motivation/pleasure but lower diminished expression were identified (and vice-versa), and these patient subgroups differed on severity of external validators such as premorbid adjustment, functional outcome, and social cognition. The demographic differences among patients statistically classified into one of those two negative symptom sub-profiles is consistent with the notion that African Americans and Caucasians differ in their relative balance of severity among these two factors, as ethnicity was to some extent differentially associated with the negative symptom profiles. Alternatively, the findings may reflect rater bias, and that clinicians (who are predominantly Caucasian in our clinic) have greater difficulty rating emotional expressivity in individuals from cultures that are different than their own. In fact, this

Table 16.7 Scale psychometrics: SANS—convergent and discriminant validity

	Caucasian	African American
BPRS positive symptoms	0.06	0.04
BPRS negative symptoms	0.73***	0.63***
BPRS disorganized symptoms	0.14	0.02
BPRS total symptoms	0.40***	0.31***

* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

is a well-known phenomenon called the "own-race-face-bias", whereby individuals are more accurate at perceiving emotion in individuals of their own culture than they are in other cultures (Malpass and Kravitz 1969). Perhaps this face identification bias extends here to clinical ratings, resulting in differential item associations between emotional expressivity items and motivation/pleasure items in African Americans with psychotic disorders.

2C. Convergent Validity: The SANS total score was highly correlated with the BPRS negative symptom score in Caucasians and African Americans, suggesting good convergent validity. The magnitude of the correlation coefficient between SANS total scores and BPRS negative scores was numerically higher in Caucasians than African Americans, but this difference was not statistically significant.

2D. Discriminant Validity: Comparison of the correlations among the SANS total score, and the BPRS Positive, Disorganized, and Total symptom subscales indicated good discriminant validity in Caucasians and African Americans (Table 16.4).

2E. Comparison of Mean SAPS Scores: Caucasians and African Americans did not significantly differ in severity on any of the four SANS global items (all p's > 0.42) (see Tables 16.2, 16.7).

Specific Recommendations for Using the SAPS and SANS with African American Clients Overall, the results of our psychometric analyses indicated that the SAPS and SANS demonstrated good reliability in terms of internal consistency in Caucasians and African Americans. Individual items included within the SAPS and SANS seem to validly measure a single latent construct, as the scales were intended, in both ethnic groups. Furthermore, the SAPS and SANS each demonstrated good convergent validity in relation to the BPRS in Caucasians and African Americans, suggesting that these scales show strong relationships with another scale purported to assess similar constructs. However, the SAPS demonstrated better convergent validity with disorganization in African Americans than Caucasians, although the correlations were sufficiently high in both cases to indicate that the SAPS disorganization items have good convergent validity in both groups.

Factor analytic results on the SAPS and the SANS were interesting, and indicated different factor structures in Caucasians and African Americans. It is possible that these differences in factor structure reflect rater bias in evaluating positive and negative symptoms. Further research is needed on this matter to determine whether rater bias might be at play on the SAPS and SANS; however, we suspect that it might given the widely documented effects of rater bias on diagnosis reviewed earlier in this chapter. Notably, such differences in factor structure emerged in the absence of absolute differences in symptom severity, suggesting that any rater bias that is present may influence how similar clinicians see different symptom dimensions, rather than the global level of psychopathology. Much like with diagnosis, the evaluation of symptom severity on popular psychiatric rating scales might also be affected by rater bias and how clinicians cluster symptoms together in people of different ethnicities when making ratings. In general, the results of our psychometric analyses indicate that two very widely used measures, the SAPS and SANS, demonstrate comparable psychometric properties in Caucasians and African Americans meeting criteria for psychotic disorders. Reliability and validity estimates, at least in terms of the analyses that were conducted here, indicate that these scales are adequate for use in African

Americans with schizophrenia; however, it is still possible that these measures are subject to some of the same rater biases that occur when making diagnoses, and this should be explored in future studies.

Conclusions

In conclusion, studies have consistently indicated an increased prevalence of schizophrenia diagnoses in African American clients. Diagnostic errors, biases in ratings, and lack of reliance on structured diagnostic procedures may contribute to this diagnostic finding, along with other sociocultural influences (e.g., urban density, socioeconomic status). Clinicians diagnosing African American individuals presenting with psychosis would be better advised to utilize structured diagnostic tools like the SCID (First et al. 2002), which has been found to reduce rater bias and increase the accuracy of diagnostic procedures. Furthermore, clinicians should consider the cultural relevance of positive, negative, and disorganized symptoms when evaluating African American clients in clinical and research settings. Failure to do so could result in inaccurate diagnoses that are based upon insufficient information, which tend to see distinct symptom dimensions as more similar than they are (i.e., positive, negative, and disorganized), thereby making over-diagnosis more likely and impeding the clinician's ability to make finely tuned treatment recommendations. To make diagnostic judgments more accurate, clinicians should consider the client's own interpretation of their symptoms in relation to cultural context, as some symptoms reflect culturally acceptable manifestations of distress.

Few studies have examined cultural differences in symptom ratings made using standard psychiatric rating scales used to measure the positive, negative, and disorganized symptoms of schizophrenia. The data presented in this chapter suggest that two of the most common scales, the SAPS and SANS, are reliable and valid for use in African Americans with psychotic disorders. The items and anchors on these scales do not appear to require modification for use in African Americans. Given their sound psychometric properties in African American individuals, clinicians could consider using these instruments to supplement other diagnostic instruments like the SCID, which tend to be less detailed in their coverage of individual positive and negative symptoms. However, these rating scales may not be immune to the rater bias problems that are thought to affect formal diagnostic procedures, and clinicians should take cultural context into consideration when using formal rating scales to assess symptom severity. Negative symptom assessments in particular should take into account cultural factors influencing normative emotional expressivity and quantity of speech, and how these may differ by culture in relation to changes in everyday context.

References

Adebimpe, V. R. (1981). Overview: White norms and psychiatric diagnosis of black patients. *American Journal of Psychiatry, 183,* 279–285.

Adebimpe, V. R., Chu, C. C., Klein, H. E., & Lange, M. H. (1982). Racial and geographic differences in the psychopathology of schizophrenia. *American Journal of Psychiatry, 139,* 888–891.

Alphs, L., Morlock, R., Coon, C., Cazorla, P., Szegedi, A., & Panagides, J. (2011). Validation of a 4-item negative symptom assessment (NSA-4): A short, practical clinical tool for the assessment of negative symptoms in schizophrenia. *International Journal of Methods in Psychiatric Research, 20,* 31–37.

Andreasen, N. C. (1983). *Scale for the assessment of negative symptoms (SANS)*. Iowa City: University of Iowa.

Andreasen, N. C. (1984). *Scale for the assessment of positive symptoms (SAPS)*. Iowa City: University of Iowa.

Andreasen, N. C. (1989). The scale for the assessment of negative symptoms (SANS): Conceptual and theoretical foundations. *British Journal of Psychiatry, 7,* 49–58.

Andreasen, N. C., & Flaum, M. (1991). Schizophrenia: The characteristic symptoms. *Schizophrenia Bulletin, 17,* 27–49.

Andreasen, N. C., Arndt, S., Alliger, R., Miller, D., & Flaum, M. (1995). Symptoms of schizophrenia: Methods, meanings, and mechanisms. *Archives of General Psychiatry, 52,* 341–351.

Arnold, L. M., Keck, P. E., Collins, J., Wilson, R., Fleck, D. E., Corey, K. B., Amicone, J., Adebimpe, V. R., & Strakowski, S. M. (2004). Ethnicity and first-rank symptoms in patients with psychosis. *Schizophrenia Research, 67,* 207–212.

Arseneault, L., Cannon, M., Fisher, H. L., Polanczyk, G., Moffitt, T. E., & Caspi, A. (2010). Childhood trauma and children's emerging psychotic symptoms: A genetically sensitive longitudinal cohort study. *American Journal of Psychiatry, 168,* 65–72.

Axelrod, B. N., Goldman, R. S., & Alphs, L. D. (1993). Validation of the 16-item negative symptom assessment. *Journal of Psychiatric Research, 27,* 253–258.

Barrio, C., Yamada, A. M., Atuel, H., Hough, R. L., Yee, S., Berthot, B., & Russo, P. A. (2003). A tri-ethnic examination of symptom expression on the positive and negative syndrome scale in schizophrenia spectrum disorders. *Schizophrenia Research, 60,* 259–269.

Blanchard, J. J., & Cohen, A. S. (2006). The structure of negative symptoms within schizophrenia: Implications for assessment. *Schizophrenia Bulletin, 32,* 238–245.

Bourque, F., Van der Ven, E., & Malla, A. (2011). A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. *Psychological Medicine, 41,* 897–910.

Bresnahan, M., Begg, M. D., Brown, A., Schaefer, C., Sohler, N., Insel, B., Vella, L., & Susser, E. (2007). Race and risk of schizophrenia in a US birth cohort: Another example of health disparity? *International Journal of Epidemiology, 36,* 751–758.

Cantor-Graae, E., & Selten, J. P. (2005). Schizophrenia and migration: A meta-analysis and review. *American Journal of Psychiatry, 162,* 12–24.

Cantor-Graae, E., & Pedersen, C. B. (2007). Risk of schizophrenia in second-generation immigrants a Danish population based cohort study. *Psychological Medicine, 37,* 485–494.

Carpenter, W. T., Strauss, J. S., & Muleh, S. (1973). Are there pathognomonic symptoms in schizophrenia. *Archives of General Psychiatry, 28,* 847–852.

Corcoran, C., Perrin, M., Harlap, S., Deutsch, L., Fennig, S., Manor, O., Nahon, D., Kimhy, D., Malaspina, D., & Susser, E. (2009). Effect of socioeconomic status and parents' education at birth on risk of schizophrenia in offspring. *Social Psychiatry and Psychiatric Epidemiology, 44,* 265–271.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). Structured clinical interview for DSM-IV-TR Axis I disorders, Research Version, Patient Edition (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute, November 2002.

Gara, M. A., Vega, W. A., Arndt, S., Escamilla, M., Fleck, D. E., Lawson, W. B., Lesser, I., Neighbors, H. W., Wilson, D. R., Arnold, L. M., & Strakowski, S. M. (2012). Influence of patient race and ethnicity on clinical assessment in patients with affective disorders. *Archives of General Psychiatry, 69,* 593–600.

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: The psychotic symptom rating scales (PSYRATS). *Psychological Medicine, 29,* 879–889.

Jablensky, A., Sartorius, N., Ernberg, G., Anker, M., Korten, A., Cooper, J. E., Day, R., & Bertelsen, A. (1992). Schizophrenia: Manifestations, incidence and course in different cultures. A World Health Organization ten-country study. *Psychological Medicine, 20,* 1–97.

Jenkins, R., Bhugra, D., Bebbington, P., Bruha, T., Farrell, M., Coid, J., Fryers, T., Weich, S., Singleton, N., & Meltzer, H. (2008). Debt, income and mental disorder in the general population. *Psychological Medicine, 38,* 1485–1493.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13,* 261–276.

Keefe, R. S., Harvey, P. D., Lenzenweger, M. F., Davidson, M., Apter, S. H., Schmeidler, J., Mohs, R. C., & Davis, K. L. (1992). Empirical assessment of the factorial structure of clinical symptoms in schizophrenia: Negative symptoms. *Psychiatry Research, 44,* 153–165.

Keller, W. R., Fischer, B. A., & Carpenter, W. T. (2011). Revisiting the diagnosis of schizophrenia: Where have we been and where are we going? *CNS Neuroscience and Therapeutics, 17,* 83–88.

Kelley, M. E., van Kammen, D. P., & Allen, D. N. (1999). Empirical validation of primary negative symptoms: Independence from effects of medication and psychosis. *American Journal of Psychiatry, 156,* 406–411.

Kirkpatrick, B., Buchanan, R. W., McKenney, P. D., Alphs, L. D., Carpenter, W. T. Jr.(1989). The schedule for the deficit syndrome: An instrument for research in schizophrenia. *Psychiatry Research, 30,* 119–123.

Kirkpatrick, B., Strauss, G. P., Nguyen, L., Fischer, B. F., Daniel, D., Cienfuegos, A., & Marder, S. R. (2011). The brief negative symptom scale: Psychometric properties. *Schizophrenia Bulletin, 37,* 300–305.

Kring, A. M., Gur, R. E., Blanchard, J. J., Horan, W. P., & Reise, S. P. (2013). The clinical assessment interview for negative symptoms (CAINS): Final development and validation. *The American Journal of Psychiatry, 170,* 165–172.

Lawson, W. (1986). Racial and ethnic factors in psychiatric research. *Culture, Medicine, and Psychiatry, 37,* 50–54.

Lipton, A. A., & Simon, F. S. (1985). Psychiatric diagnosis in a state hospital: Manhattan State revisited. *Hospital and Community Psychiatry, 36,* 368–373.

Liss, J., Welner, A., Robins, E., & Richardson, M. (1973). Psychiatric symptoms in white and black inpatients I: Record study. *Comprehensive Psychiatry, 14,* 475–481.

Lukoff, D., Nuechterlein, K. H., & Ventura, J. (1986). Manual for the expanded brief psychiatric rating scale. *Schizophrenia Bulletin, 13,* 261–276.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13,* 330–334.

March, D., Hatch, S. L., Morgan, C., Hatch, S. L., Morgan, C., Kirkbride, J. B., Bresnahan, M., Fearon, P., & Susser, E. (2008). Psychosis and place. *Epidemiologic Review, 30,* 84–100.

Minsky, S., Vega, W., Miskimen, T., Gara, M., & Escobar, J. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. *Archives of General Psychiatry, 60,* 637–644.

Mueser, K. T., Sayers, S. L., Schooler, N. R., Mance, R. M., & Haas, G. L. (1994). A multisite investigation of the reliability of the scale for the assessment of negative symptoms. *American Journal of Psychiatry, 151,* 1453–1462.

Mukherjee, S., Shukla, S., Woodle, J., Rosen, A. M., & Olarte, S. (1983). Misdiagnosis of schizophrenia in bipolar patients: A multiethnic comparison. *American Journal of Psychiatry, 140,* 1571–1574.

Murray, R. M., & Hutchinson, G. (1999). *Psychosis in migrants: The striking example of African-Caribbeans resident in England, in search for the causes of schizophrenia, Volume 4: Balance of the Century* (pp. 129–140). Darmstadt: Steinkopff.

Myers, N. L. (2011). Update: Schizophrenia across cultures. *Current Psychiatry Reports, 13,* 305–311.

Neighbors, H. W., Jackson, J. S., Campbell, L., & Williams, D. (1989). Racial influences on psychiatric diagnosis: A review and suggestions for research. *Community Mental Health Journal, 25,* 301–311.

Neighbors, H. W., Trierweiler, S. J., Munday, C., Thompson, E. E., Jackson, J. S., Binion, V. J., & Gomez, J. (1999). Psychiatric diagnosis of African Americans: Diagnostic divergence in clinician structured and semistructured interviewing conditions. *Journal of the National Medical Association, 91,* 601–612.

Neighbors, H. W., Trierweiler, S. J., Ford, B., & Muroff, J. R. (2003). Racial differences in diagnosis using a semi-structured instrument: The importance of clinical judgment. *Journal of Health and Social Behavior, 44,* 237–256.

Ödegaard, Ö (1932). Emigration and insanity. *Acta Psychiatrica et Neurologica Scandinavica, 14,* 1–206.

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological Reports, 10,* 149–165.

Overall, J. E., & Gorham, D. R. (1988). The brief psychiatric rating scale (BPRS): Recent developments in ascertainment and scaling. *Psychopharmacology Bulletin, 24,* 97–99.

Peralta, V., & Cuesta, M. J. (1995). Negative symptoms in schizophrenia: A confirmatory factor analysis of competing models. *American Journal of Psychiatry, 152,* 1450–1457.

Sayers, S. L., Curran, P. J., & Mueser, K. T. (1996). Factor structure and construct validity of the scale for the assessment of negative symptoms. *Psychological Assessment, 8,* 269–280.

Schneider, K. (1959). *Clinical psychopathology*. (trans: Hamilton, M. W.). New York: Grune and Stratton.

Schofield, P., Ashworth, M., & Jones, R. (2010). Ethnic isolation and psychosis: re-examining the ethnic density effect. *Psychological Medicine 2010, FirstView,* 1–7.

Selten, J. P., & Cantor-Graae, E. (2005). Social defeat: Risk factor for schizophrenia? *British Journal of Psychiatry, 187,* 101–102.

Selten, J. P., & Cantor-Graae, E. (2007). Hypothesis: Social defeat is a risk factor for schizophrenia? *British Journal of Psychiatry, 191,* 9–12.

Sharpley, M. S., Hutchinson, G., McKenzie, K., & Murray, R. M. (2001). Understanding the excess of psychosis among the African-Caribbean population in England. *British Journal of Psychiatry, 178,* S60–S68.

Simon, R. J., Fleiss, J. L., Gurland, B. J., Stiller, P. R., & Sharpe, L. (1973). Depression and schizophrenia in hospitalized black and white mental patients. *Archives of General Psychiatry, 28,* 509–512.

Strakowski, S. M., Shelton, R. C., & Kolbrener, M. L. (1993). The effects of race and comorbidity on clinical diagnosis in patients with psychosis. *Journal of Clinical Psychiatry, 54,* 96–102.

Strakowski, S. M., Flaum, F., Amador, X., Bracha, H. S., Pandurangi, A. K., Robinson, D., & Tohen, M. (1996a). Racial differences in the diagnosis of psychosis. *Schizophrenia Research, 21,* 117–124.

Strakowski, S. M., McElroy, S. L., Keck, P. E. Jr., & West, S. A. (1996b). Racial influence on diagnosis in psychotic mania. *Journal of Affective Disorders, 39,* 157–162.

Strauss, G. P., Hong, L. E., Keller, W. R., Buchanan, R. W., Gold, J. M., Fischer, B. A., McMahon, R. P., Catalano, L. T., Culbreth, A. J., Carpenter, W. T., & Kirkpatrick, B. (2012). Factor structure of the brief negative symptom scale. *Schizophrenia Research, 142,* 96–98.

Strauss, G. P., Horan, W. P., Kirkpatrick, B., Fischer, B. A., Keller, W. R., Miski, P., Buchanan, R. W., Green, M. F., & Carpenter, W. T. (2013). Deconstructing negative symptoms of schizophrenia: Avolition-apathy and diminished expression clusters predict clinical presentation and functional outcome. *Journal of Psychiatric Research, 47,* 783–790.

Trierweiler, S. J., Neighbors, H. W., Munday, C., Thompson, E. E., Jackson, J. S., & Binion, V. J. (2006). Differences in patterns of symptom attribution in diagnosing schizophrenia between African American and non-African American clinicians. *American Journal of Orthopsychiatry, 76,* 154–160.

Worthington, C. (1992). An examination of factors influencing the diagnosis and treatment of black patients in the mental health system. *Archives of Psychiatric Nursing, 6,* 195–204.

Assessing Somatoform Disorders with the African American Client

17

Samantha M. Behbahani and Robert M. Greenstone

The somatoform disorders (SFD's) represent a defined group of mental disorders, which are distinguished by chronic conditions. They consist of multiple medically unexplained bodily complaints that occur over a prolonged period of time (Hiller and Janca 2003). Although the symptoms may have varying presentations based on cultural influences, the concept of SFD's has been accepted amongst clinicians worldwide. They were originally introduced in 1980 by the third Diagnostic and Statistical Manual of Mental Disorders (DSM-III) as an encompassing term covering various clinical conditions, such as Somatization Disorder, Conversion Disorder, Psychogenic Pain Disorder, and Hypochondriasis. The current classification systems, DSM-IV-TR (2000), and International Statistical Classification of Diseases and Related Health Problems (ICD-10), have continued to use definitions very similar to those originally introduced in DSM-III, but have added Body Dysmorphic Disorder and Undifferentiated Somatoform Disorder. Pain Disorder was substituted for Psychogenic Pain Disorder. Due to this common origin, the diagnostic frameworks for SFD's in DSM-IV-TR (2000) American Psychiatric Association (2000) and ICD-10 (1992) are nearly identical and a broad international consensus on the classification of SFDs has been established. Another reason for the similarity between the DSM-IV-TR and the ICD-10 is related to the fact that both classifications focus on symtomatology instead of etiology.

When any group of mental disorders is identified to be distinctive from others, a number of assessment tools and instruments are created to help to identify them in research and clinical settings, and SFD's are no exception. In different countries, researchers suggested self- or observer-rated methods that can be used to detect somatoform patients, as well as explain their clinical features and measure longitudinal changes (Hiller and Janca 2003).Some instruments were thoroughly evaluated using psychometric and statistical methods. As a result, a number of those with good psychometric quality were used in different studies and gave rise to cross-cultural comparisons.

Pain is the most commonly reported presenting complaint in medical settings (Hardin 1998). The importance of appreciating the level and quality of patients' pain has recently come to the forefront with the implementation of recommendations for the addition of pain level as a vital sign (American Pain Society Quality of Care Committee 1995). The importance of acknowledging ethnic differences in patient report of pain has been cited over three decades (Zatzick and Dimsdale 1990). This chapter represents an attempt to provide an overview of most thoroughly evaluated and widely used assessment

---

S. M. Behbahani (✉) · R. M. Greenstone
Psychology Department, Carlos Albizu University, 2173 N.W. 99th Avenue, Miami, FL 33172, USA
e-mail: sbehbahani@albizu.edu

R. M. Greenstone
e-mail: rgreenstone541@sunmail.albizu.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_17,

instruments in the field of the SFDs in the African American (AA) population.

Somatic symptoms are an associated feature of anxiety disorders that have received little research attention among non-white samples. In addition, most previous studies have examined the effects of somatic symptoms in a cross-sectional instead of prospective manner (Kingery et al. 2007). Furthermore, it is not possible to assume the relationship between somatic complaints and psychosocial functioning. Evidence suggests that the prevalence of anxiety in AA youth is high and that these youth may have a significant tendency to express this anxiety in the form of physical symptoms (Kingery et al. 2007). Anxiety sensitivity (i.e., the belief that anxiety-related symptoms will have harmful effects) was previously measured in a sample of urban AA elementary school children and results indicated that they endorsed higher levels of anxiety sensitivity than their white counterparts (Lambert et al. 2004). In addition, AA adults have been found to express symptoms of anxiety and depression in the form of somatic symptoms (Cooper-Patrick et al. 1999; Neal and Turner 1991; Robins and Regier 1991; Snowden and Pingitore 2002).

Assessment of Somatoform Disorders

Broadly used classification systems, such as the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) and the International Classification of Disease (ICD-10) World Health Organization, employ comparable symptom clusters to describe somatoform disorders (SFD's, i.e., somatization, undifferentiated somatoform, conversion, pain, hypochondriasis, and body dysmorphic). However, the instruments available for SFD's are not a homogenous group. They vary according to their chief diagnostic purpose of assessment. Furthermore, many mental health professionals agree that the current characterizations of SFD's do not capture the most frequently noted symptoms of clients with somatic complaints (Escobar and Gureje 2007). The DSM-V American Psychiatric Association (2010) recognizes the significant overlap found between the current SFD's (e.g., Somatization, Hypochondriasis, Undifferentiated Somatoform, and Pain) and intends to reclassify them in to a *complex somatic symptom disorder* (http://www.dsm5.org/proposedrevision/Pages/SomaticSymptomDisorders.aspx).

Assessment of Somatoform Disorders in AAs

Strategies, methodologies, techniques, and psychological instruments are challenged to establish their applicability for the population on which they are used. The assessment of somatoform disorders in AAs is complicated by the nonexistence of appropriate psychological tests and fact that AAs are a heterogeneous group. This poses the question as to whether race-specific norms should be created for existing instruments. Many psychologists and researchers have been critical of the use of psychological assessments administered to AAs. Appropriate standards to be assessed for AAs must take into account characteristics such as perceptions of racism, responses to racism, acculturation, identity development and formation, and self-consciousness (Lindsey 1998).

A number of barriers have been linked to mental health-seeking behaviors. These can include, but are not limited to, “stigma, lack of resources such as poverty, inadequate coverage, transportation, and childcare” (LaQuita and Hatcher 2012). These types of challenges often lead to great disparities in minority populations, including AAs. LaQuita and Hatcher (2012) noted other barriers that include insufficient understanding of the mental health profession, institutional racism, discrimination, and cultural mistrust of mental health professionals and the medical establishment. Many individuals in the black community believe that only God will prevent and cure mental illness, that problems will cease if ignored, and that involving oneself in mental health treatment signifies weakness and diminished pride (Thompson et al. 2004). This perceived stigma is so great that people often refuse to seek much needed help. Unfortunately, there are many potentially severe consequences for not locating therapeutic intervention. For AAs and other racial and ethnic minorities, the family unit provides the primary support for individuals with mental health problems. However, social barriers can arise against relatives and household members of the stigmatized person. The behaviors of persons with severe psychological disorders may isolate the family, damage its reputation, and damage relationships with community members (Lefley 1989).

Despite the lack of research studies that examine somatic symptoms among AA youth, one study (White and Farrell 2006) examined two specific somatic symptoms (headaches and abdominal pain) among a primarily AA adolescent sample ($N=528$; age range 11–14 years). Findings of this study demonstrated that a sizeable portion of the sample experienced headaches (40 %) and abdominal pain (36 %) on at least a weekly basis. In addition, AA and European American patients appear to report significantly different levels and qualities of pain in both laboratory and clinical conditions (Breitbart et al. 1996; Edwards et al. 2001; Edwards and Fillingim 1999; Faucett et al. 1994; Stewart et al. 1996).

Some research suggests that African Americans tend to report depression in the form of somatic complaints, whereas Caucasian Americans are more inclined to report depression in a cognitive-affective manner (Blazer et al. 1998; Greko et al. 1996; Simon et al. 1973). Furthermore, it appears as though AA patients report lower pain tolerances despite similar pain thresholds. In the Epidemiologic Catchment Area studies, somatization disorder was found in 0.01 % of the population and was most prevalent among AA women (0.8 %) followed by AA men (0.4 %) (Cassisi et al. 2004). These differences point to cultural factors rather than biological factors and considerations must be made for assessments.

Diagnostic Instruments

The Schedules for the Clinical Assessment in Neuropsychiatry

The Schedules for Clinical Assessment in Neuropsychiatry (SCAN) is a semi-structured clinician directed interview created under the framework of the WHO/Alcohol,Drug Abuse, and Mental Health Administration (ADAMHA) (Wing et al. 1990).The SCAN was created to integrate both the objective facts in an interview, yet also include the variable information presented by the patients.All information assessed in the interview is formulated to provide a diagnosis, which can be coded as both DSM-IV or ICD-10 diagnoses (Sartorius and Janca 1996). Administration time can be lengthy, averaging approximately 90 min to completion time (Sartorius and Janca 1996). SFD's assessment section is relatively short, yet addresses health-related concerns and physical complaints (Hiller and Janca 2003). Validity and reliability factors are not available for the AA population, yet field trials indicate good reliability in 20 centers in 14 countries (Wing et al. 1990).

The Composite International Diagnostic Interview

The Composite International Diagnostic Interview (CIDI) is a lay-administered assessment tool, comprised of 276 structured symptom-based questions, followed by probing questions (Wittchen 1994). Clinical experience is not required to administer this measure, and it can be completed in 90 min. The reliability of the CIDI has been widely studied and has proven to be an excellent indicator of psychiatric diagnoses. However, the research on somatoform disorder assessment with the CIDI has been sparse (Hiller and Janca 2003), due to the limitations a somatizing sample in psychiatric settings. Wittchen (1994) found test–retest reliability for somatization disorder, 0.74, pain disorder, 0.68 and 0.71 hypochondriasis. Furthermore, Janca et al. (1995) administered the CIDI within two settings, outpatient psychiatric clinics and primary care centers, and found test–retest reliability was 0.76 for somatoform disorders as a group. Inter-rater reliability for the somatization questions was good, at 0.67 (Wittchen et al. 1991).

Limitations of the CIDI have been the lengthy administration time. The range of completion time varies from 78 min (experienced administrators) to 103 min (novice administrators) (Hiller and Janca 2003). Furthermore, somatoform disorder-oriented questions include extensive probing, leading to a lengthier interview, comparable to the questions related to substance abuse disorders and mood disorders. Lastly, validation studies within the AA population is limited Kessler and Bedirhan Ustun (2004).

Screening Instruments

The General Health Questionnaire

The General Health Questionnaire (GHQ) is a self-report, screening questionnaire for identifying psychiatric disorders (GHQ-60; Goldberg and Hiller 1979). The original version was comprised of 60 items, but abbreviated versions have been created thereafter (GHQ-30, GHQ-28, GHQ-12). In epidemiological studies, a 12-items version would be most beneficial to utilize due to its shorter length. The reported Cronbach alpha coefficient for the GHQ ranges between 0.82 and 0.86.

Bogner (2004) measured psychological distress and urinary incontinence (UI) between community-dwelling older AAs and Caucasian Americans ($n=747$; adults were at least 50 years of age). Psychological distress was assessed using the General Health Questionnaire (GHQ). Results indicated that AAs with UI were more likely to experience psychological distress as measured using the GHQ than were AAs without UI (unadjusted odds ratio =4.22, 95% confidence interval =1.72–10.39) while the association between UI and psychological distress did not achieve statistical significance in Caucasian Americans (Bogner 2004). These results point to the tendency for AAs be more greatly affected by UI than for Caucasian Americans, however, a more substantial correlation between medical diagnoses and psychological distress must be explored through future research.

Instruments of Severity Ratings

The Minnesota Multiphasic Personality Inventory-2

The Minnesota Multiphasic Personality Inventory-2 is a widely known objective personality measure, assessing a variety of traits indicative of psychopathology. The MMPI-2 has been revised once since its first version, with one goal to restandardize the test with a new sample of adults, to more accurately establish new test norms. A total of 2600 adults were utilized in the sample, with 1138 males and 1462 females assessed. The MMPI-2 was restandardized to more accurately display the ethnicity distribution of the USA. Out of the 1138 males assessed, 11.1% were AA, and 12.9% of the females assessed were AA (Butcher et al. 1990a). These percentages were chosen to reflect the 1990 US Census. Furthermore, upon analysis of the most current US Census (census.gov), the AA population is approximately 13.6% of the population. Therefore, the most recent standardization is fairly comparable to the current census.

The MMPI-2 has two clinical scales that can be utilized to assess for somatoform symptomatology, scale 1-hypochondriasis (Hs), and scale 3-hysteria (Hy). Scale 1-hypochondriasis, assesses vague and unrealistic somatic concerns, which are often not validated with biomedical markers (Hiller and Janca 2003). The Hysteria scale evaluates the individual's lack of awareness and vulnerabilities, often in reaction to stressful situations (Hiller and Janca 2003). In addition, there is a one content scale, Health Concerns (HEA), which assesses for gastrointestinal, cardiovascular, neurological, sensory, dermatological, respiratory, and pain complaints (Butcher et al. 1990b). When assessing the possibility of somatic diagnoses, the aforementioned scales are often evaluated to determine the severity of health related and medically unexplained symptomatology.

When utilizing the MMPI-2 to assess possible somatoform disorders in the AA client, a review of the literature within this population is warranted. There have been a number of studies performed to investigate the differences AA patients may illustrate on those scales highlighting somatic concerns. Timbrook and Graham (1994) utilized part of the MMPI-2 restandardization sample, and thus analyzed the results for any ethnic differences on the validity and clinical scales for both males and females. Results illustrated that on scales 1 (Hy) and 3 (Hs) failed to reveal clinically significant differences between the AA and Caucasian sample of men and women, with fairly small effect sizes (0.31, 0.22, respectively). The validity of the MMPI-2 was also assessed by McNulty et al. (1997), whereby 123 AA outpatients were compared against 561 Caucasian outpatients. Results failed to find significant differences on scale 1 (Hs), scale 3 (Hy), or the Health Concerns scale (HEA) between the AA and Caucasian population.

A meta-analytic review was conducted on 37 studies comparing AA and Caucasian samples utilizing the MMPI and MMPI-2 (Nagayama Hall et al. 1999). For the male sample, small effect sizes (−0.52–0.33) revealed that AAs exhibited higher scores on scale 1 (Hs). Caucasians displayed higher scores on scale 3 (Hs), yet also displaying small effect sizes (−0.42–0.25). When analyzing the female AA versus female Caucasian sample, similar results were found on scales 1 and 3, with small effect sizes (Nagayama Hall et al. 1999). Furthermore, Arbisi et al. (2002) investigated differences between 229 AA and 1558 Caucasian psychiatric inpatients. Results indicated that AA men and women did not differ from their Caucasian counterparts on any clinical scale measuring somatic concerns. However, AA men scored significantly higher than Caucasian men on the Health Concerns (HEA) content scale ($p<0.001$).

Millon Adolescent Clinical Inventory (MACI)

The Millon Adolescent Clinical Inventory (MACI) supplements the Millon Adolescent Personality Inventory (MAPI). It was specifically developed for use in clinical, residential, and correctional settings. It is particularly useful in the evaluation of troubled adolescents, and may be used for diagnostic assistance, in formulating treatment plans, and as an outcome measure (Millon and Davis 1993). The MACI consists of a total of 31 scales: 12 Personality Patterns scales (Axis II), eight Expressed Concerns Scales, seven Clinical Syndrome Scales, three Modifying Indices (which assess particular response styles), and a Validity scale (Millon and Davis 1993).

Barry and Grilo (2002) used the MACI (Millon and Davis 1993) to assess ethnic differences between adolescent Caucasians and AAs in regard to body image disturbances. The body image disturbance factor differed significantly by ethnicity ($p<0.001$) as did the MACI-Body Disapproval score ($p<0.005$). Scheffe post-hoc tests illustrated that Caucasian adolescents had significantly higher raw scores on the MACI-Body Disapproval than did AAs (mean difference = 2.76, 95 % confidence interval [CI] = 0.17–5.35, $p<0.05$, two-tailed test). Similarly, Scheffe post-hoc test demonstrated that Caucasian adolescents had significantly higher body image disturbance than their AA counterparts (mean difference = 0.47, 95 % CI = 0.00–0.76, $p<0.05$, two-tailed test) Millon (2000).

The Symptom Checklist-90-Revised

The Symptom Checklist-90-Revised (SCL-90-R; Derogatis 1994) is comprised of 90 items, grouped into nine symptom clusters that generate distress. These clusters encapsulate symptoms of somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism. The measure is designed to screen for a wide range of psychological illnesses. The reliability of all SCL-90-R subscales is satisfactory, and validity measures have established a high degree of convergent and concurrent validity of the subscales (Derogatis 1994).

Ayalon and Young (2009) were the first to evaluate the appropriateness of the use of the Symptom Checklist-90-Revised (SCL-90-R) with AA college students. The sample consisted of 66 Caucasian American and 70 AA college students in a large Midwestern city. Differences in symptomatology reports, both related and unrelated to the subscale pathology, were assessed. Reliability of SCL-90-R scales was high in both groups and ranged from 0.77 to 0.89 for AAs (median = 0.86) and from 0.78 to 0.92 for Caucasian Americans (median = 0.85). Overall, findings indicated that there were few group differences between AAs and Caucasian Americans on the SCL-90-R (Ayalon and Young 2009). There were disparities between the two racial groups on the Phobic Anxiety subscale, which is suggestive of a higher degree of phobic anxiety AAs due to external circumstances. Findings support the use of the SCL-90-R with AA college students (Ayalon and Young 2009).

The Whitely Index

The Whiteley Index (WI) is a questionnaire that was developed to identify hypochondriacs. This particular test attempts to distinguish among subgroups of hypochondriacs, namely those who are preoccupied with symptoms, those who fear developing a serious disease, and those who are convinced they are already ill (Pilowsky 1967). Pilowsky's original statistical analyses demonstrated that only 8 of the 14 items have a factor loading of 0.40 or greater on any target factor. Research that attempted to replicate Pilowsky's findings of a three-factor model has come up short, which leaves the characterization of the measure's factor structure in question (Fink et al. 1999; Speckens et al. 1996). This suggests either flawed conceptualization of the hypochondriacal factors or poor item sampling (Schmidt 1994).

There is a consensus that the Whiteley Index only assesses a single general hypochondriacal fears and convictions dimension (Speckens et al. 1996). Speckens et al. (1996) produced a single factor solution based on 10 items. The internal consistency ($\alpha = 0.76$ to $\alpha = 0.80$) was borderline across three diverse samples, and the mean interitem correlations were found to be in an acceptable range (rs = 0.25–0.26). As a single scale, the Whiteley Index emerge as one that is unidimensional. Although the WI has proven its validity and reliability when assessing for global hypochondriacal characteristics, standardized norms for the AA population are not yet available. Overall, there has been limited systematic progress towards understanding hypochondriasis as a dimensional construct. First, there is a lack of agreement on how the most important factors of the hypochondriasis construct should be classified. Second, there is no complete, empirically validated model that includes all key dimensions. Third, the measures currently in existence have limited construct validity.

The Illness Attitude Scales

The Illness Attitudes Scale (IAS) is a 29-item measure that evaluates fears, attitudes and beliefs associated with hypochondriacal concerns and abnormal illness behavior (Kellner et al. 1987). It is a self-rated measure comprised of nine subscales (worry about illness, concern about pain, health habits, hypochondriacal beliefs, thanatophobia, disease phobia, bodily preoccupation, treatment experience, and effects of symptoms). Kellner's goal of the IAS was to construct a valid tool, which would adequately discriminate the attitudes and beliefs found in patients with Hypochondriasis while limiting the amount of symptom overlap. Kellner et al. (1987) reports that the IAS appears to be a valid and reliable measure for identifying hypochondriacal patients. The reported Cronbach's reliability coefficients of the scale ranged from 0.062 to 1.00.

One validation study compared 21 patients diagnosed by the DSM-III with Hypochondriasis to an equal number of psychiatric patients, family practice patients, and a group of employees/non-psychiatric subjects with paralleled demographics. Kellner concluded that the IAS is "best used as a state measure rather than a trait measure" (Kellner 1987). Although standardized norms of the IAS for the AA demographic are not available, it is crucial to consider the importance of longstanding patterns of hypochondriacal concerns that are greatly influenced by social constructs, such as feeling marginalized in a Whitedominated society. Unfortunately, the IAS fails to appreciate these complexities and it should not be used as an equivalent replacement for a clinical interview.

The Multidimentional Inventory of Hypochondriacial Traits

The Multidimensional Inventory of Hypochondriacal Traits (MIHT, Longley et al. 2005) is a relatively new instrument utilized to address the limitations of previous measures of hypochondriasis, such as the Whitely Index and the Ilness Attitude Scales.This instrument is a self-report measure that assesses four domains of hypochondriasis: Affective, Cognitive, Perceptual, and Behavioral (Stewart et al. 2008). More specifically, affective refers to the individual's anxiety regarding their health, cognitive pertaining to their beliefs about threat of illness, perceptual illustrating one's responsiveness to physical symptoms, and lastly, behavioral addressing one's medical care seeking to quell their health anxiety.

Validity and reliability of the MIHT is fairly strong. The validity measures displayed an internal consistency of 0.80 for all four scales. Furthermore, test–restest reliability was also good, ranging between 0.75 and 0.78. Furthermore, when the MIHT was compared to other measures assessing hypochondriasis, convergent validity was promising. For example, when assessing convergence with the Whitely Index two scales were utilized from the MIHT, Cognitive and Affective. Convergent analysis displayed good coefficient alpha (0.84 and 0.87). Furthermore, coefficient alpha was also good when the MIHT was assessed against the Illness Attitude Scales/Health Anxiety Questionnaire, particularly in the domains of Affective, Cognitive, and Perceptual dimensions (0.83–0.89). Results were not as promising when assessing the Behavioral dimension against the Illness Attitude Scales/Health Anxiety Questionnaire (0.57; Stewart et al. 2008).

The MIHT is a 31-item, five-point like rt-scale test. It was normed on a combination of university students, community members, and medical outpatients. More than half of the demographics of the university students (N=1673) is unknown, as they responded anonymously. The members of the community (N=172) consisted mainly of Caucasian participants (95%), with only one participant being AA (0.6%). The medical outpatient sample consisted of 120 participants, with the majority being Caucasian (89.2%; Longley et al. 2005). Given the most recent US Census, this population is poorly reflective of the presentation of the AA population, and thus may be a poor measure to analyze the presentation of hypochondriasis in this population. Furthermore, according to the authors' most recent literature review, there appears to be a void of research addressing the utility of the MIHT in the AA population.

Brief Symptom Inventory

The Brief Symptom Inventory (BSI) is a 53-item self-report measure based on a likert scale. This measure is derived from the SCL-90, assessing the same areas of distress. The BSI assesses nine areas of psychological distress, with subscales including 'somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism" (Derogatis 2000). This assessment measure is fairly rapid in administration, with an average completion time of 8–10 min. Within the subscale of Somatization (SOM), internal consistency was strong (0.80), yet a lowered test–retest reliability was displayed (0.68). Furthermore, the correlation coefficient (0.93) on the Somatization scale was consistent with the SCL-90, confirming the test's validity. The BSI was normed on four separate sample groups: adult psychiatric outpatients, adult nonpatients, adult psychiatric inpatients, and adolescent nonpatients. A total of 4807 patients were assessed with the measure, yet the representation of AAs in this sample equaled approximately only 27% of the total sample (Derogatis 2000).

In a comparison of ethnically diverse female clients, authors found AA and Mexican-American patients scored higher on all scales of the BSI, including the Somatic scale. Based on this study, results are consistent with much of the literature discussing the concept that AA women display higher rates of somatization than the Caucasian counterparts (Golding and Karno 1988; Hemmings et al. 1998). Hemmings et al. (1998) also reported that the AA women in the aforementioned study endorsed more symptomatology indicating higher scores on the Positive Symptom Total, resulting from greater emotion and anxiety-based responses.

McGill Pain Questionnaire—Long and Short Forms

The McGill Pain Questionnaire was developed to specifically address the sensory, affective, and qualitative views of pain, from the patient's perspective (Katz and Melzack 2011). Due to the varied experience of pain, intensity is also assessed. The McGill Pain Questionnaire was developed in 1975, and since its development the short form version has emerged, McGill Pain Questionnaire-Short Form (MPQ-SF) Melzack (1987). The original version consists of 78 pain descriptors that provides an overall index of pain. The MPQ-SF consists of 15 items measuring two domains of pain, sensory and adaptive. Additionally, this measure includes an embedded Visual Analog Scale (VAS), utilized to assess the intensity of the pain experienced. The VAS has been proven to be responsive to changes in the pain stimulus. The MPQ-SF correlates highly with the major indices of the long-form McGill Pain Questionnaire (0.67–0.94).Internal consistency for the MPQ-SF estimates are 0.9 = 84 for the affective domain, and 0.076 for the sensory domain (Katz et al. 2011). Lastly, the SF-MPQ-2 was developed to specifically address neuropathic pain, which was limited in the previous versions of the test (Katz et al. 2011). 'The SF-MPQ-2 has very good to excellent psychometric properties, including adequate to high internal consistency reliability estimates for the subscale (0.73–0.87) and total scores (0.91–0.95)' (Katz et al. p. 53).

Cassisi et al. (2004) examined ethnic differences in the use of pain descriptors among AA and European American patients. The Short-Form McGill Pain Questionnaire, including the embedded VAS, was utilized to assess chronic pain syndromes from a total of 489 participants, 97 identifying as AA, and 392 identifying as European American. Results indicated that there may be systematic differences in how the pain descriptors are used by each group. The AA sample tended to utilize more physically based symptoms, such as gnawing, throbbing, and punishing. The Caucasian sample utilized both physically-based and affectively-charged descriptions, such as stabbing, fearful and heavy. A drawback of this study is the relatively small sample size of the AA group, compared to the European American group.

Baker and Green (2005) assessed intra-race differences among older (>50) and younger (<50) AA patients utilizing the McGill Pain Inventory, compared to younger and older European Americans. Within group results of AA reported higher MPQ scores for the younger cohort, yet there were no actual differences related to the experience of suffering or the disability attached to the pain. Furthermore, analysis revealed that older AAs exhibited better coping skills in regard to pain, when compared to the younger sample (3.1 +/− 1.6 versus 3.8 +/− 1.5, $p < 0.01$). Similar results were found for the European American sample.

The Short Health Anxiety Inventory

The Short Health Anxiety Inventory (SHAI; Salkovskis et al. 2002) is an 18-item, shortened and revised version of the 64-item Health Anxiety Inventory, designed to capture three domains of hypochondriasis, including worry about health, awareness of bodily sensations, and feared consequences of having an illness. The SHAI was conceptualized from a cognitive-behavioral perspective Fergus and Valentiner (2011). The SHAI is reported to have high internal consistency, acceptable Croncach's alpha scores, as well as strong construct validity. Wheaton et al. (2010) assessed the utility of the SHAI in a student sample, whereby 12.1 % of the sample included AA men and women. Internal consistency of the study was good, at 0.88. Furthermore, Abramowitz et al. (2007) assessed the psychometric properties and construct validity in a sample of university students. Within this study, AA participants consisted of 18.8 % of the sample. Upon analysis of ethnic differences on the SHAI, results indicated that AAs, Asian Americans, and Caucasians did not differ from each other, yet Hispanic Americans displayed significantly lower SHAI scores. A recent meta-analysis of the SHAI evaluated 78 papers that cited the use of the aforementioned measure (Alberts et al. 2013).This systematic review reported that the SHAI is a psychometrically sound measure, yet there is a dearth of research evaluating the usage among diverse samples of patients.

Body Image Disturbances Among AAs

Body image disturbance is an increasing issue in Western cultures and is correlated with numerous mental health problems including anorexia, bulimia, body dysmorphia, and depression (Pimenta et al. 2009). *The aesthetic expectancies and pressures found in Western societies, based on thinness for women and a muscular physique for men, are thought to have a great influence on body image disturbance* (Friedman et al. 2002; Keel et al. 2001). Most women wish to lose weight, even when their physical parameters are in the normative range (Muennig et al. 2008). Similarly, a significant number of men have reported muscle dissatisfaction (Cafri and Thompson 2004). Body Dysmorphic Disorder (BDD) is a psychological disorder that consists of a distressing or impairing preoccupation with an imagined or slight defect in appearance (Didie et al. 2006). Individuals with BDD have impaired psychosocial functioning, markedly poor quality of life, and high rates of suicidal ideation and attempts.

The way one perceives his or her body can have reverberating effects on his or her self-esteem. North American culture, which is predominantly Caucasian American-oriented, places additional pressures on AAs whose natural body types differ from societal standards. With the amount of media that depicts the epitome of beauty as a Caucasian American male or female, body image disturbance among AAs can quickly take place. These challenging physical expectancies are much greater to fulfill, which can result in internalized distress Ruffalo et al. (2006).

Yale-Brown Obsessive Compulsive Scale

The Yale-Brown Obsessive Compulsive Scale (Y-BOCS) is a “reliable, valid, and widely-used scale that is the standard measure of obsessive-compulsive disorder (OCD) severity” (Phillips et al. 1997). In order to address the severity of Body Dysmorphic Disorder (BDD), Phillips et al. (1997) developed the BDD-YBOCS, a 12-item semi-structured clinician-rated instrument. It aims to assess obsessional preoccupation with a perceived defect of one’s physical appearance, functional impairment, and distress level of the patient. The measure demonstrated high interrater reliability and acceptable test–retest reliability (intraclass r for total score = 0.88). Cronbach’s alpha coefficient was 0.80, which is indicative of sufficient scale homogeneity.

Total score on the BDD-YBOCS was significantly correlated with universal measures of symptom severity: it was positively correlated with the Clinical Global Impressions Scale (CGI) score ($r=0.55$, $p=0.003$) and was negatively correlated with the Global Assessment of Functioning (GAF) score ($r=-0.51$, $p<0.001$). The BDD-YBOCS seems to be a reliable assessment of BDD and proves adequate validity. Phillips et al. (1997) used a sample of 63 men and 62 women (mean age 32.8 + 9.5 years), but did not describe the demographic characteristics of their sample, so the Y-BOCS has yet to demonstrate its applicability with the AA population.

The Body Image Rating Scale (BIRS)

The Body Image Rating Scale (BIRS) is used to both screen and assess severity of Body Dysmorphic Disorder (BDD). It measures affective, cognitive, and behavioral associations of BDD. The scale's internal validity was reported at coefficient $\alpha = 0.93$ and test–retest reliability of $r = 0.86$. Convergent validity with the Body Dysmorphic Disorder Examination Self-Report scale was $r = 0.86$ (Mayville et al. 1998). In general AAs were found to be more satisfied (F (6, 450) = 4.06, $p = 0.0006$) with their bodies than all other ethnic groups (Mayville et al. 1999).

Measures

Assessment	Type of assessment
BIRS	Body image
BSI	General psychiatric severity index
CIDI	Diagnostic interview
GHQ	Screening tool
IAS	Hypochondriasis; severity index
MACI	Adolescent/general psychiatric
MIHT	Hypochondriasis; severity index
MMPI-2	General psychiatric; hypochondriasis, hysteria, health concerns
MPQ (Long, Short, SF-2)	Pain; severity index
SCAN	Diagnostic interview
SCL-90	General psychiatric; severity index
SHAI	Hypochondriasis; severity index
WI	Hypochondriasis; severity index
YBOCS-BDD	Body image

BIRS Body Image Rating Scale, *BSI* Brief Symptom Inventory, *CIDI* Composite International Diagnostic Interview, *GHQ* General Health Questionnaire, *IAS* Illness Attitude Scales, *MACI* Millon Adolescent Clinical Inventory, *MMPI-2* Minnesota Multiphasic Personality Inventory-2, *MPQ* McGill Pain Questionnaire, *SCAN* Schedules for Clinical Assessment in Neuropsychiatry, *SCL-90* Symptom Checklist-90-Revised, *SHAI* Short Health Anxiety Inventory, *WI* Whitely Index, *YBOCS-BDD* Yale-Brown Obsessive Compulsive Scale for Body Dysmorphic Disorder

Summary and Recommendations

The instruments reviewed in this chapter include a variety of measures assessing different severities of somatic complaints. The diagnostic instruments include the Composite International Diagnostic Interview (CIDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). The CIDI has the ability to administer by non-professionals, yet the SCAN must be administered by a trained psychologist or psychiatrist. The SCAN and CIDI are better equipped to be utilized in research environments, and are less applicable for usage in clinical settings, such as hospitals or outpatient mental health centers.

The MMPI-2, SCL-90, Brief Symptom Inventory (BSI), and the MACI are well-validated, widely used instruments specifically utilized to assess general psychiatric states in patients. These measures include scales to assess the severity of symptoms related to somatic concerns, body complaints, and affective experiences related to the aforementioned. The MMPI-2, SCL-90, and the BSI have been normed and validated on AAs, and thus provide reliable and valid means to measure somatic concerns in this population.

Other discussed measures in this chapter were those assessments specifically created to address and evaluate circumscribed symptomatology, including hypochondriacal beliefs, pain, and body image disturbances. Although most of these measures were minimally normed on the AA population, there are a number of studies assessing their utility with this population. It is recommended to not only consider the score each client attains on these measures, but their ability to express their emotions, emote their somatic conditions, as well as being cognizant of the varying social conditions that may be present in their daily life Table 17.1.

Table 17.1 At-a-glance summary of measures

Assessment name	Disorder assessed	Method of administration	Brevity of measure (min)
BIRS	Body Dysmorphic Disorder	Clinician administration	10
BSI	Hypochondriasis; general health concerns	Self-administration	25–30
CIDI		Lay administration/Interview	90
GHQ	General health concerns	Self-report/Screening	60
IAS	Hypochondriasis; general health concerns	Self-administration	15–20
MACI	Adolescent; body concerns	Self-administration	30
MIHT	Hypochondriasis	Self-administration	15–20
MMPI-2	Hypochondriasis; Hysterical features	Self-administration	120–180
MPQ (Long, Short, SF-2)	Pain Disorder	Self-administration	5–10
SCAN	Health related concerns; physical complaints	Clinician administered/ Interview	90
SCL-90	Somatization concerns	Self-administration	15
SHAI	Hypochondriasis	Self-administration	10–15
WI	Hypochondriasis	Self-administration	5
YBOCS-BDD	Body Dysmorphic Disorder	Clinician administered	10

References

Abramowitz, J., Deacon, B., & Valentiner, D. (2007). The Short Health Anxiety Inventory: Psychometric properties and construct validity in a non-clinical sample. *Cognitive Therapy and Research, 31,* 871–883.

Alberts, N., Hadjistavropoulos, H., Jones, S., & Sharpe, D. (2013). The Short Health Anxiety Inventory: A systematic review and meta-analysis. *Journal of Anxiety Disorders, 27*(1), 68–78.

American Pain Society Quality of Care Committee. (1995). Quality improvement guidelines for the treatment of acute pain and cancer. *Journal of the American Medical Association, 274,* 23.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). doi:10.1176/appi.books.9780890423349.

American Psychiatric Association. (2010). DSM-5 development. http://www.dsm5.org/proposedrevision/pages/somaticsymptomdisorders.aspx.

Arbisi, P., McNulty, J., & Ben-Porath, Y. (2002). A comparison of the MMPI-2 validityin African American and Caucasian psychiatric in patients. *Psychological Assessment, 14*(1), 3–15.

Ayalon, L., & Young, M. (2009). Using the SCL-90-R to assess distress in African Americans and Caucasian Americans. *Journal of Black Studies, 39*(3), 420–433.

Baker, T., & Green, C. R. (2005). Intrarace differences among black and white Americans presenting for chronic pain management: The influence of age, physical health, and psychosocial factors. *Pain Medicine, 6*(1), 29–38.

Barry, D., & Grilo, C. (2002). Eating and body image disturbances in adolescent psychiatric inpatients: Gender and ethnicity patterns. *Wiley Periodicals, Inc, 32,* 335–343.

Blazer, D., Landerman, L., Hays, J., Simonsick, E., & Saunders, W. (1998). Symptoms of depression among community-dwelling elderly African Americans and Whiteolder adults. *Psychological Medicine, 28,* 1311–1320.

Bogner, H. (2004). Urinary incontinence and psychological distress in community-dwelling older African Americans and whites. *Journal of the American Geriatrics Society, 52*(11), 1870–1874.

Breitbart, W., McDonald, M., Rosenfeld, B., Passik, S., Hewitt, D., Thaler, H., & Portenoy, R. (1996). Pain and ambulatory AIDS patients: I. Pain characteristics and medical correlates. *Pain, 68,* 315–321.

Butcher, J. N., Dahlstrom, J. R., Graham, J. R., Tellegen, A., & Kaemmer, B. (1990a). *Minnesota Multiphasic Personality Inventory (MMPI-2): Manual foradministration and scoring*. Minneapolis: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Williams, C. L., & Ben-Porth, Y. S. (1990b). *Developmentand use of the MMPI-2 content scales*. Minneapolis: University of Minnesota Press.

Cafri, G., & Thompson, J. (2004). Evaluating the convergence of muscle appearance attitude measures. *Assessment, 11,* 224–229.

Cassisi, J., Umeda, M., Deisinger, J., Sheffer, C., Lofland, K., & Jackson, C. (2004). Patterns of pain descriptor usage in African Americans and European Americans with chronic pain. *Cultural Diversity and Ethnic Minority Psychology, 10*(1), 81–89.

Cooper-Patrick, L., Gallo, J., Powe, N., Steinwachs, D., Eaton, W., & Ford, D. (1999). Mental health service utilization by African Americans and Whites: The Baltimore epidemiologic catchment area follow-up. *Medical Care, 37,* 1034–1045.

Derogatis, L. R. (1994). *Manual for the symptom checklist 90-revised*. Minneapolis: Pearson Assessments.

Derogatis, L. R. (2000). *Brief Symptom Inventory (BSI): Administration, scoring, and procedures manual*. Minneapolis: NCS Pearson.

Didie, E., Tortolani, C., Walters, M., Menard, W., Fay, C., & Phillips, K. (2006). Socialfunctioning in body dysmorphic disorder: Assessment considerations. *Psychiatry Quarterly, 77,* 223–229.

Edwards, R., Doleys, D., Fillingim, R., & Lowery, D. (2001). Ethnic differences in pain tolerance: Clinical implications in a chronic pain population. *Psychosomatic Medicine, 63,* 316–323.

Edwards, R., & Fillingim, R. (1999). Ethnic differences in thermal pain responses. *Psychosomatic Medicine, 61,* 346–354.

Escobar, J., & Gureje, O. (2007). Influence of cultural and social factors on the epidemiology of idiopathic somatic complaints and syndromes. *Psychosomatic Medicine, 69*(9), 841–845.

Faucett, J., Gordon, N., & Levine, J. (1994). Differences in postoperative pain severity among four ethnic groups. *Journal of Pain and Symptom Management, 9,* 383–389.

Fergus, T., & Valentiner, D. (2011). The short health anxiety inventory and multidimensional inventory of hypochondriacal traits: A comparison of two self-report measures of health anxiety. *Cognitive Therapy and Research, 35,* 566–574.

Fink, P., Ewald, H., Jensen, J., Sorensen, L., Engberg, M., Holm, M., & Munk-Jorgensen, P. (1999). Screening for somatization and hypochondriasis in primary care and neurological in-patients: a seven-item scale for hypochondriasis and somatization. *Journal of Psychosomatic Research, 46*(3), 261–273.

Friedman, K., Reichmann, S., Costanzo, P., & Musante, G. (2002). Body image partially mediates the relationship between obesity and psychological distress. *Obesity Research, 10,* 33–41.

Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine, 9*(1), 139–145.

Golding, J. M., & Karno, M. (1988). Gender differences in depressive symptoms among Mexican Americans and Non-Hispanic Whites. *Hispanic Journal of Behavioral Sciences, 10*(1), 1–19.

Greko, P., Brikman, A., & Routh, D. (1996). Depression and coping in candidates for kidney transplantation: Racial and ethnic differences. *Journal of Clinical Psychology in Medical Settings, 3,* 337–353.

Hall, Gordon C. Nagayama Bansal, Anita

Hardin, K. (1998). Chronic pain management. In P. M. Camic & S. J. Knight (Eds.), *Clinical handbook of health psychology: A practical guide to effective interventions* (pp. 123–164). Kirkland: Hogrefe & Huber.

Hemmings, M., Madrigal, D., Reimman, J., & Velasquez, R. J. (1998). Predictors of scores on the Brief Symptom Inventory for ethnically diverse female clients. *Psychological Reports, 83,* 800–802.

Hiller, W., & Janca, A. (2003). Assessment of somatoform disorders: A review of strategies and instruments. *Acta Neuropsychiatrica, 15,* 167–179.

Janca, A., Burke, Jr., Isaac, M., Burke, K. C., Costa, J. A., Silva, E., Acuda, S. W., Altamura, A. C., Chandrashekar, C. R., Miranda, C. T., & Tacchini, G. (1995). Original article: The world health organization somatoform disorders schedule. A preliminary report on design and reliability. *In European Psychiatry, 10*(8), 373–378.

Katz, J., & Melzack, R. (2011). The McGill Pain Questionnaire: Development, psychometric properties, and usefulness of the long form, short form, and short-form-2. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 45–66). New York: Guilford.

Keel, P., Mitchell, J., Davis, T., & Crow, S. (2001). Relationship between depression and body dissatisfaction in women diagnosed with bulimia nervosa. *International Journal of Eating Disorders, 30,* 48–56.

Kellner, R., Abbott, P., Winslow, W., & Pathak, D. (1987). Fears, beliefs, and attitudesin DSM-III hypochondriasis. *The Journal of Nervous and Mental Disease, 174,* 20–25.

Kessler, R., & Bedirhan Ustun, T. (2004). The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 93–121.

Kingery, J., Ginsburg, G., & Alfano, C. (2007). Somatic symptoms and anxiety among African American adolescents. *Journal of Black Psychology, 3*(4), 363–378.

Lambert, S., Cooley, M., Campbell, K., Benoit, M., & Stansbury, R. (2004). Assessing anxiety sensitivity in inner-city African American children: Psychometric properties of the childhood anxiety sensitivity index. *Journal of Clinical Child and Adolescent Psychology, 33,* 248–259.

LaQuita, S., & Hatcher, L. (2012). African Americans are less likely to seek mental health treatment. *HIV Clinician, 24*(4), 11.

Lefley, H. (1989). Family burden and family stigma in major mental illness. *American Psychologist, 44,* 556–560.

Lindsey, M. (1998). Culturally competent assessment of African American clients. *Journal of Personality Assessment, 70*(1), 43–53.

Longley, S., Watson, D., & Noyes, R. (2005). Assessment of the hypochondriasis domain: The Multidimensional inventory of Hypochondriacal Traits (MIHT). *Psychological Assessment, 17*(1), 3–14.

Mayville, S., Gipson, M., & Katz, R. (1998). *Body image rating scale*. Poster presented at the annual meeting of the Western Psychological Association, Albuquerque, New Mexico.

Mayville, S., Katz, R., Gipson, M., & Cabral, K. (1999). Assessing the prevalence of body dysmorphic disorder in an ethnically diverse group of adolescents. *Journal of Child and Family Studies, 8,* 357–362.

McNulty, J., Graham, J., Ben-Porath, Y., & Stein, L. (1997). Comparative validity of the MPPI-2 scores of African American and Caucasian mental health center clients. *Psychological Assessment, 9*(4), 464–470.

Melzack, R. (1987). The short-form McGill Pain Questionnaire. *Pain, 30*(2), 191–197.

Millon (2000). The Millon Adolescent Clinical Inventory and psychopathy. *Journal of Personality Assessment, 75*(1), 110–125.

Millon, T. Davis, R. D. (1993). The Millon Adolescent Personality Inventory and the Millon Adolescent Clinical Inventory. *Journal of Counseling & Development, 71*(5), 570–574.

Muennig, P., Jia, H., Lee, R., & Lubetkin, E. (2008). I think therefore I am: Perceived ideal weight as a determinant of health. *American Journal of Public Health, 98,* 501–506.

Nagayama Hall, G., Bansal, A., & Lopez, I. (1999). Ethnicity and psychopathology: Ameta-analytic review of 31 years of comparative MMPI/MMPI-2 research. *Psychological Assessment, 11*(2), 186–197.

Neal, A., & Turner, S. (1991). Anxiety disorders research with African Americans: Current status. *Psychological Bulletin, 109,* 400–410.

Phillips, K., Hollander, E., Rasmussen, S., Aronowitz, B., DeCaria, C., & Goodman, W. (1997). A severity rating scale for body dysmorphic disorder: Development, reliability, and validity of a modified version of the Yale-Brown obsessive compulsive scale. *Psychopharmacology Bulletin, 33*(1), 17.

Pilowsky, I. (1967). Dimensions of hypochondriasis. *British Journal of Psychiatry, 113,* 89–93.

Pimenta, A., Sanchez-Villegas, A., Bes-Rastrollo, M., Lopez, C., & Martinez-Gonzalez, M. (2009). Relationship between body image disturbance and incidence of depression: The SUN prospective cohort. *BMC public health, 9,* 1.

Robins, L., & Regier, D. (1991). *Psychiatric disorders in America*. New York: Free.

Ruffalo, J., Phillips, K., Menard, W., Fay, C., & Weisberg, R. (2006). Comorbidity of body dysmorphic disorder and eating disorders: Severity of psychopathology and body image disturbance. *International Journal of Eating Disorders, 39*(1), 11–19.

Salkovskis, P. M., Rimes, K. A., Warwick, H. M. C., & Clark, D. M. (2002). The Health Anxiety Inventory: Development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychological Medicine, 32*(5), 843–853.

Sartorius, N., & Janca, A. (1996). Psychiatric assessment instruments developed by the World Health Organization. *Society of Psychiatry and Psychiatric Epidemiology, 31,* 55–69.

Simon, R., Fleiss, J., Gurland, B., Stiller, P., & Sharpe, L. (1973). Depression and schizophrenia in hospitalized Black and White mental patients. *Archives of General Psychiatry, 28,* 509–512.

Snowden, L., & Pingitore, D. (2002). Frequency and scope of mental health service delivery to African Americans in primary care. *Mental Health Services Research, 4,* 123–130.

Speckens, A., Spinhover, P., Sloekers, P., Bolk, J., & van Hemert, A. (1996). A validation study of the Whiteley Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practicepatients. *Journal of Psychosomatic Research, 40*(1), 95–104.

Stewart, W., Lipton, R., & Liberman, J. (1996). Variation in migraine prevalence by race. *Neurology, 47,* 52–59.

Stewart, S., Sherry, S., Watt, M., Grant, V., & Hadjistavropoulos, H. (2008). Psychometric evaluation of the Multidimensional Inventory of Hypochondriacial Traits: Factor structure and relationship to anxiety sensitivity. *Journal of Cognitive Psychotherapy: An International Quarterly, 22*(2), 97–114.

Timbrook, R. E., & Graham, J. R. (1994). Ethnic differences on the MMPI-2? *Psychological Assessment, 6*(3), 212.

Thompson, V., Bazile, A., & Akbar, M. (2004). African Americans' perceptions of psychotherapy and psychotherapists. *Professional Psychology: Research and Practice, 35,* 19–26.

Wheaton, M., Berman, N., Franklin, J., & Abramowitz, J. (2010). Health anxiety: Latent structure and associations with anxiety-related psychological processes in a student sample. *Journal of Psychopathology and Behavioral Assessment, 32,* 565–574.

White, K., & Farrell, A. (2006). Anxiety and psychosocial stress as predictors of headache and abdominal pain in urban early adolescents. *Journal of Pediatric Psychology, 31,* 582–596.

Wing, J. K., Babor, T., Brugha, T., Burke, J., Cooper, J. E., Giel, R., Jablenski, A., Regier, D., & Sartorius, N. (1990). SCAN: Schedules for Clinical Assessment in Neuropsychiatry. *Archives of General Psychiatry, 47,* 589–593.

Wittchen, H. U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research, 28*(1), 57–84.

Wittchen, H. U., Robins, L. N., Cottler, L. B., Sartorius, N., Burke, J. D., Regier, D., & participants in the multicentre WHO/ADAMHA field trials (1991). Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). *The British Journal of Psychiatry, 159,* 645–653.

World Health Organization. (2008). ICD-10: International statistical classification of diseases and related health problems (10th Rev. ed.). New York, NY.

Zatzick, D., & Dimsdale, J. (1990). Cultural variations in response to painful stimuli. *Psychosomatic Medicine, 52,* 544–557.

18

School-based Assessment with African American Children and Adolescents

Tara C. Raines, Jennifer Twyford and Erin Dowdy

School-Based Assessment with African American Children and Adolescents

In the last two decades, there has been increasing pressure on schools to demonstrate student progress using school-based assessment data. Legislation such as No Child Left Behind (NCLB) and revisions to the Individuals with Disabilities in Education Act (IDEA) promote the use of data to assist teaching and learning. School-based assessments are also used to measure individual achievement and evaluate the effectiveness of specific interventions and programs (Reddy et al. 2009; Van Der Heyden and Burns 2005). This movement toward data-based educational accountability encourages school districts and local school administrators to actively monitor student response to instruction (RtI) and interventions using a combination of standardized, district-level and classroom level assessments (Fuchs and Fuchs 2006; Jimerson et al. 2007).

While controversial, the implications of student performance on school-based assessments are tremendous. Outcomes from school-based assessments are used to inform decisions regarding school performance grades, a metric often referenced in school choice debates (Hastings and Weinstein 2008). These grades, in some school districts, could dictate whether or not parents are mandated to enroll their students in a particular school or have the option to choose another. These data can also be linked to teacher evaluations, incentives, or retention (Baker et al. 2010; Borman and Kimball 2005).

School-based assessments also have the potential to lead to pernicious student outcomes. As used for special education placement, traditional school-based assessment practices have historically over-identified African American students for disability services and under-identified them for gifted education programs. Specifically, African American students are overrepresented in the most restrictive educational environments, underrepresented in the least restrictive environments, and overrepresented in the special education categories of emotional disturbance, mild mental retardation, moderate mental retardation, learning disabilities, and speech and language impairments (Skiba et al. 2006). Although many factors have been identified as contributing to

---

T. C. Raines (✉)
Department of Educational Psychology and Higher Education, University of Nevada Las Vegas,
4505 S. Maryland Parkway, P.O. Box 453003,
Las Vegas, NV 89154, USA
e-mail: Tara.Raines@unlv.edu

J. Twyford
Graduate School of Psychology, California Lutheran University, 2201 Outlet Center Drive,
Suite 600, Oxnard, CA 93036, USA
e-mail: jtwyford@callutheran.edu

E. Dowdy
Department of Counseling, Clinical, and School Psychology, University of California Santa Barbara,
Bldg. 275, Room 4100, Santa Barbara,
CA 93106, USA
e-mail: edowdy@education.ucsb.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_18,

disproportionality, including group differences in performance on tests of academic achievement (Hosp and Reschley 2004), poor performance, and testing bias (Ysseldyke et al. 2000 as cited in Fore et al. 2006), the resulting outcomes for African American students are poor.

Failure to identify African American students for gifted programs can lead to missed opportunities for rigorous and specialized instruction aimed to cultivate student strengths and provide them with advanced academic opportunities (Cross and Donovan 2002; Ford and Harmon 2001). The disproportionate placement of African American students in special education programs is even more troubling as research has demonstrated that students identified for special education services may be twice as likely to drop out of high school, receive substandard education, and suffer social isolation and low self esteem (Harris et al. 2004; Terras et al. 2009; Thurlow et al. 2002; Waitoller et al. 2010). Specifically in regards to dropout rates, studies have found there to be profound social and economic implications for the students who do not complete high school, their families, and society. Students who have not completed high school have high rates of unemployment, make less money, are more likely to need public assistance, and are more likely to become involved with the criminal justice system (Ferri and Connor 2005). It could be concluded that the practice of identifying a disproportionate number of African American students for special education services and overlooking them for gifted programs places them on a trajectory for diminished life opportunities.

Fortunately, there have been advances in the design and implementation of school-based assessments that can be particularly beneficial for African American students, reducing disproportionality and promoting inclusion in programs for high achieving students. School-based assessments, when implemented according to current best practices, can provide the flexibility of using a student or population-based metric for assessing achievement. Whereas, traditional assessment practices often compare students to groups outside of their community who may or may not have access to the same rigor of instruction or support, more current approaches consider the instructional or behavioral context, as well as student culture and experience when interpreting assessment data. Similarly, school-based assessment can provide the framework for early intervention and graduated levels of support services that focus intervention resources on students who are in the most need.

This chapter will discuss the current use of school-based assessments for evaluating student academic performance and behavioral–emotional functioning. Specifically, this chapter will provide information on the application of a population-based continuum for assessing social, emotional, behavioral, and academic functioning that includes screening, intervention, and progress monitoring before diagnostic decisions are made. Practices for measuring academic achievement including curriculum-based measures and norm-referenced approaches will be examined and their use with African American students will be discussed. Finally, specific considerations for the use of culturally-responsive school-based assessment practices with African American students and barriers to optimal implementation are reviewed.

A Population-Based Assessment Model

Traditionally, school-based assessment has been primarily concerned with determining eligibility for special education. However, more current and comprehensive models of school psychological assessment begin in the general education setting and provide the opportunity for the assessment of all individuals with a focus on prevention, early intervention, and reducing biases (Lau and Blatchley 2009). Specifically, schools have been moving away from individual, reactive approaches of assessment towards population-based, prevention approaches (Harrison et al. 2004).

Population-based service delivery models, inclusive of assessment for academic, behavioral, and emotional needs, ensure that the needs of all children are met (Doll and Cummings 2008). Population-based models frequently have multi-tiered levels of assessment and interventions with tiers that vary in intensity. For example, within a three-tiered model, *universal* assessments are provided to all children within a given population (e.g., school, grade, classroom), *selected* assessments are provided to children who have already been identified as at-risk, and *indicated* assessments are provided to students with the greatest level of impairment or need (Durlak 1997). School-based assessment for social, emotional, behavioral, and academic functioning occurs at each of these levels with the goal of integrating general, remedial, and special education services for improved outcomes for all students (Reschly 2008). The following sections will describe the application of population-based approach for the school-based assessment of emotional and behavioral problems.

Universal Assessments

Educators have long recognized the importance of screening as an essential first step within a population-based approach (Glover and Albers 2007). Screening data can be utilized to provide information to identify and monitor the needs of an entire population (Dowdy et al. 2010). For example, based on a school-wide screening for behavioral and emotional risk, administrators may be able to identify key problematic areas, such as anxiety problems in 9th graders who are new to the school. Similarly, data can be analyzed across ethnicities or gender to determine if African American males are exhibiting more externalizing behavior problems or if they are simply being suspended at higher rates. Screening data can then be used to make data-based decisions regarding what types of interventions are needed, and for whom, and then can subsequently be used as a means to evaluate student progress towards the stated school intervention goals.

In addition to using screening data to make population-based decisions, screening can also be used to identify specific students who may benefit from prevention or early intervention activities. The basic premise behind screening is that by engaging in a time-efficient process to identify students with potential risk or early signs of difficulty, supports can be put into place to assist students. Screening for risk is differentiated from screening for disorders, with emphasis placed on identifying symptoms early in their trajectory when prevention activities are more likely to be beneficial (Kamphaus 2012).

Schools are currently engaged in screening for a variety of different problems, including screening for hearing, vision, academic, and behavioral problems (Vannest 2012). The assessment methods vary widely depending on the nature of the screening, with some schools utilizing existing data, such as suspension or office disciplinary referrals, to identify the students that may be at risk of behavioral or emotional problems. However, some schools are engaged in a more pro-active and preventative screening process, such as requiring all students to complete a brief behavioral rating scale to assess for potential symptoms of risk or difficulty (Chin et al. 2013).

Universal screening for behavioral and emotional risk, in particular, is an important component of a culturally-competent assessment practice, as it provides all students the same opportunity for potential identification and service provision (Dowdy et al. 2013). As the outcomes associated with special education placement are discouraging (Waitoller et al. 2010), recent research has investigated if universal screening for behavioral and emotional risk can reduce disproportionality rates (Dever et al. 2013). The presumption is that by changing the primary mechanism through which students are referred for special education, the students who ultimately receive special education services may differ. Although this line of research is in its infancy, initial findings are promising.

Historically, referral to special education for emotional or behavioral difficulties occurs through a teacher nomination process whereby teachers note their concerns with students who are then tested for, and often placed into, special education (Dowdy et al. 2011b; Gerber and Semmel 1984). This reliance on subjective teacher referrals presents with a number of concerns, with research demonstrating that referral decisions may differ when teachers use more objective and structured means of identifying students (Eklund et al. 2009). Similarly, the reliance on existing data, such as office discipline referrals or suspensions as a means of determining who is referred for special education, is fraught with problems. Decades of discipline disproportionality research has provided evidence that minority students are more likely to be suspended, sent out of class, or expelled for school misbehavior (Skiba et al. 2002). This suggests that relying on existing data, as a means of universal screening, would only exacerbate existing problems of minority overrepresentation in certain special education categories. Alternatively, a systematic universal screening process for emotional and behavioral concerns, may allow for a more data-driven and empirical method for determining who may benefit from additional assessment or intervention.

When evaluating screening assessments for use in schools, Glover and Albers (2007) recommend examining the assessment's appropriateness for use with the intended population, technical characteristics, as well as their usability within a specific context. Assessment measures should be evaluated across various populations, including ethnic and minority groups, to determine measurement equivalence (Tyson 2004). Modern behavioral assessment screening measures are often normed on populations representative of the general population of U.S. children with regard to sex, race/ethnicity, and clinical or special education classification (Reynolds and Kamphaus 2004). However, there is limited psychometric evidence available for screeners and their use with diverse subpopulations (Severson et al. 2007), representing a current research need in the field. Among available screening instruments that have been found to have strong psychometric properties for use with culturally and linguistically diverse populations are the Behavior Assessment System for Children: Behavioral and Emotional Screening System (BASC-2: BESS; Kamphaus and Reynolds 2007) and the Strengths and Difficulties Questionnaire (SDQ; Goodman and Goodman 1999). Findings of various studies suggest measurement equivalence across racial/ethnic groups (Bourdon et al. 2005; DiStefano and Morgan 2010; Dowdy et al. 2011a; Raines et al. 2014; Ravens-Sieberer et al. 2008).

Selected Assessments

Following a universal screening indicative of increased risk, a more comprehensive assessment process may ensue. The more in-depth assessment should not be designed primarily to answer eligibility questions, but rather to gather additional information regarding the student's functioning and how to best intervene. For example, following administration of the teacher-rated BASC-2: BESS (Kamphaus and Reynolds 2007), a student may receive an extremely elevated *T* score suggestive of behavioral and emotional risk. As a screener is designed solely to provide an overall level of risk, a school-based practitioner would likely be interested in learning what specific types of risk the student is exhibiting so that interventions could be tailored accordingly. An omnibus teacher rating scale, such as the BASC-2 Teacher Rating Scale (Reynolds and Kamphaus 2004), may be provided as an indicated assessment. Results could provide detailed information suggesting that the student is displaying externalizing conduct problems and hyperactivity. Practitioners can then utilize this information to design a series of behavior modification interventions.

Rating scales are only one of many potential sources of information that may be used as a selected assessment. School-based practitioners are also advised to: conduct interviews with parents, teachers, and the student; gather an extensive history of educational, school, and community functioning; conduct observations, and provide an assessment of the classroom or school context (Lau and Blatchley 2009). Parent interviews are a critical component to reducing bias when evaluating a student for potential special education placement into the category of emotional disorder (ED) or emotional or behavioral disorder (EBD) (Lau and Blatchley 2009). Information about familial and community networks and resources can be gathered as a means of understanding the student's functioning as well as providing information about potential sources of support and intervention (Palacios and Trivedi 2009). Interviews can be used to gather important contextual information impacting the student's values, behavioral standards, and identify the student's cultural values and beliefs (Lau and Blatchley 2009). These cultural values and beliefs can then be easily incorporated into the collaborative process of setting goals and treatment planning (Jones 2009).

Indicated Assessments

Following a lack of response to the interventions and continued interference with academic progress, an indicated, comprehensive assessment would follow to determine the need or suitability for additional services and interventions. This indicated assessment would include information from various informants, assess functioning within a wide variety of domains, and use multiple strategies and tools to gather comprehensive information about the problem behaviors. A multidimensional, multitask approach is imperative to get a thorough understanding of the student's presenting problems and to begin the process of generating ideas that will be useful for intervention planning (Lau and Blatchley 2009).

A final step in the assessment of emotional or behavioral problems would be consideration for special education eligibility for ED or EBD. Federal and state guidelines will need to be addressed and prior to special education placement school-based practitioners often consider child psychopathology, behavioral–emotional interactions, functional relations between the environment and the problem behaviors, and an evaluation of the effectiveness of interventions that have been implemented for the student (Lau and Blatchley 2009). Caution should be exercised when considering a placement of ED or EBD for an African American student due to the known overrepresentation of minority students in these special education categories (Skiba et al. 2002).

A review of specific assessment tools that may be used in a selected or indicated assessment is provided in various other chapters within this book (e.g., see Chap. 10 for discussion of assessment tools used for behavioral disorders). However, application within a school-based framework may differ, as the ultimate goal is to use assessments as a foundation for engaging a problem-solving model. Specifically, school-based assessments are used to determine: if there is a problem and what it is; why the problem may be occurring; what interventions can be used to help the problem; and if the interventions were effective (Tilly 2008). The following are best practices considerations for nondiscriminatory assessment within a problem-solving framework: assess for the purpose of intervention as opposed to assessing to determine eligibility; use authentic and alternative assessment procedures to complement any standardized, norm-referenced tests given; assess the learning ecology to determine if there are any extrinsic factors impacting the student's difficulties; determine if a student has had an adequate opportunity to learn thoroughly an assessment of the instructional setting, curriculum, school policies, and personnel; evaluate educationally relevant cultural factors; work to reduce biases in traditional assessment practices; and draw conclusions from data that has been integrated across multiple indicators (Ortiz 2008).

Academic Assessment

A population-based assessment framework can be similarly effective for the prevention and intervention of academic problems. Universal, selected, and indicated assessments are employed to increase the cultural appropriateness of academic achievement assessment for all students. This section discusses school-based academic assessment procedures that have been found to reduce some of the problems inherent in more traditional reactive assessment procedures. Research as it relates specifically to the assessment of African American children and adolescents is presented below.

Curriculum-Based Measurement

Curriculum-based measurement (CBM) utilizes a standard set of procedures to assess a student's progress over time on measures derived from the school's curriculum and at the student's individual instructional- or grade-level (Deno 2003). Standardized procedures inherent to CBM allow for increased reliability and validity to make individual and group comparisons. CBM probes can be administered for progress monitoring, instructional monitoring, or individual or group evaluations at the rate of weekly, annually, at benchmarks (i.e., Fall, Winter, Spring), or as needed. CBM is based on frequently administered, brief measures in reading, writing, spelling, and math computation (Fuchs and Deno 1991). CBM fluency measures in reading and math are the most common.

Reading In reading, a variety of measures exist that align with the developmental nature of reading skills. For example, in kindergarten and primary grades, CBM in reading may measure phoneme segmentation, letter naming, and rhyming. As reading skills become more advanced, reading comprehension may be measured by CBM-Maze, a metric in which students must select the correct word when approximately every seventh word is omitted and a fixed choice is provided (Ferguson et al. 1992). The most commonly used CBM in reading measures reading aloud or oral reading fluency (ORF). The standardized procedures for CBM-ORF measures oral reading performance with short, timed passages derived from the school's adopted curriculum. The student is asked to read aloud three different passages for 1 min while the examiner records the number of errors, and the median score is used as the measure of performance (Hosp and Hosp 2003).

Research regarding use of CBM-ORF with African American students has yielded mixed results. Emerging research has provided evidence for the appropriateness of the use of CBM-ORF with African American students (Betts et al. 2008; Hintze et al. 2002). However, others have found evidence of bias when measuring CBM-ORF across students of different ethnicities (African American and Caucasian; Kranzler et al. 1999). CBM reading measures that require reading aloud have evidence of validity and improving performance across different groups (Fewster and Macmillan 2002); however, measures for students younger than second grade, beginning readers, and different types of reading measures still require further investigation for use with diverse student populations (Wayman et al. 2007).

Math CBM in math is a measure of math computation (Thurber et al. 2002). Probes contain problems of single or mixed computation timed for 2–5 min, depending on the grade level. Similar to CBM-ORF procedures, three sets of probes are administered and the median score based on the correct number of digits is derived for the student's performance (Shinn 1989). The use of CBM in math as a screening procedure appears appropriate for identifying African American students with poor math performance (Christ et al. 2008); however, further evidence for use with African American students is still needed. In addition, further research is needed of CBM in math related to the psychometrics, broad utility, broad populations, and procedures (Christ et al. 2008).

Advantages and Disadvantages Although further research is needed for use of CBM with African American populations, there are advantages to using CBM. The inherent standardized and frequent administration of CBM probes allows for an individual's performance to be compared with his or herself, instead of an unrepresentative or inappropriate normative sample. The systematic approach of CBM allows for monitoring progress, goal setting, determining academic skills growth, and making individual educational programming decisions (Deno 2003). In addition, CBM can improve teacher judgment regarding student proficiency (Marston et al. 1984).

Similar to the identification of students with emotional or behavioral problems, teacher referrals have typically been used to identify students with learning difficulties. It has been suspected that these procedures have contributed to significant disproportional representation of African Americans in various special education categories (Gravois and Rosenfield 2006). As CBM is sensitive to relatively short-term prereferral educational interventions, teachers can document student progress and evaluate the effectiveness of the interventions before more restrictive educational programs are considered. In addition, CBM addresses content bias in standardized academic achievement by using measures based on the school's curriculum that have high evidence of criterion validity (Fuchs et al. 1988).

Although there are many benefits to using CBM to measure academic achievements with African American students, there are also disadvantages. Despite the voluminous research indicating that CBM is an effective strategy for enhancing learning for special education students (e.g., see Stecker et al. 2005), few studies have investigated the effects, if any, of cultural bias in CBM (Evans-Hampton et al. 2002). Researchers have indicated that CBM as a metric is not completely without bias, and that other factors outside of the measure, such as developmental level and grade, may contribute to differential performance across groups (Hintze et al. 2002; Hosp and Reschley 2004; Kranzler et al. 1999). Further investigations of CBM in reading as it pertains to African American students should examine situational effects. One investigation found no situational bias related to timing procedures for African American students (Evans-Hampton et al. 2002). However, further investigations are necessary to examine situational bias. For example, examinations into class versus pull-out administration settings, and the effect of the ethnicity of the examiner, are needed.

Local Norms

Schools and teachers are able to increase the usability of CBM and other academic assessments with African American students by developing peer comparison groups. Schools can develop "local norms" by sampling the students in the given comparison group (Deno 2003). Local norms can increase the appropriateness of across-student comparisons by creating a broadly representative comparison group, particularly in schools with rapidly changing student demographics, urban schools, or when the commercially available norms are not appropriate (Deno 2003).

As CBM is derived directly from the school's curriculum and standardized in administration, an individual's performance can be compared with an appropriate peer comparison group. For example, comparison peer norms could be developed for each classroom, at the school-wide grade level, or district-wide grade level. For CBM in reading, national norms are available from the National Reading Panel report (National Institute of Child Health and Human Development 2000; see also Hasbrouck and Tindal 2006). In addition, general norms for expected growth with CBM in reading for use with progress monitoring for students with learning disabilities are also available (see Deno et al. 2001). When content is deemed to be culturally appropriate and based on the local curriculum, CBM has evidence of validity for use with African American students (Green et al. 2005).

Fitting a Population-Based Assessment Approach

Similar to the above application with emotional and behavioral problems, schools may employ a three-tiered population-based assessment framework to evaluate student achievement. At the *universal level*, all students' academic skills may be screened using CBM at benchmark intervals (e.g., Fall, Winter, Spring) (Fore et al. 2006). School-wide benchmarks assessments may be used to develop local norms to make comparisons of student performance with their peers, predict academic performance on statewide achievement tests (Keller-Margulis et al. 2008), or identify students who may not be benefitting from the core curriculum (Fore et al. 2006).

At the *selected level*, students that would benefit from differentiated instruction receive targeted assessments and/or interventions with the goal to improve student performance and enhance academic skills. At this level, CBM may be used to monitor progress of some student's academic skills in response to an instructional intervention (Shinn 2007). At the *indicated level*, a student who continued to demonstrate academic skill deficits may receive an individual psychoeducational evaluation for additional supports and services, such as special education. In this process, practitioners may utilize additional standardized academic assessments to determine whether an educational disability exists.

School assessors should select measures with adequate psychometrics, and reflective of each student's individual characteristics, including race, ethnicity, language, regional location, age, and socioeconomic status. (See Chap. 2 for a more detailed discussion of individual psychological assessments.) After a student is identified for additional supports and services, CBM may aid practitioners in writing goals and objectives for Individualized Education Plans (Fore et al. 2006). Although CBM is currently used by schools for screening and placement decisions (Fewster and Macmillan 2002), further investigations for use of these practices with African American populations are needed.

Considerations and Barriers

Approaching *any* assessment with *any* population requires a set of specialized skills and caution when interpreting the outcomes and results. This is particularly true in school-based assessment of African American children and adolescents. Given the educational history of African Americans in the education system of the United States of America and the implications for school-based assessments it is imperative that practitioners approach these assessments mindfully.

The African American experience in the U.S. education system has been tumultuous. Early, postslavery efforts of African Americans seeking education were met with overt opposition (Fields-Smith 2005). Additionally, the transition to an integrated school system was also arduous and fraught with resistance from the dominant culture (Kluger 2011). As a result, many scholars believe that despite integration, this country is not fully committed to providing an "equal and excellent" public education to African American students (Fennimore 1997, p. 245; Hilliard 1991; Zirkel 2005). These beliefs are supported by repeated acknowledgment of the perilous educational practices and outcomes for African American students (i.e., disproportionate number of African American students in special education, underrepresentation of African American students in gifted programs, the achievement gap, etc.) coupled with very few available, viable solutions aiming to rectify this malpractice (Raines et al. 2012). Therefore, it is crucial that we acknowledge that school-based assessments can be perceived as a vehicle to promote segregated and unequal educational practices (Ferri and Connor 2005). Considering this, the development, implementation, and interpretation of school-based assessments with African American students should be approached with the primary goal of improving and enhancing their educational experience. This can be done by devoting specific attention to identifying a clear purpose for the assessment and selecting instruments that provide authentic information about student functioning.

School-based assessments are unlike most other assessments or evaluations as neither the individual being evaluated nor their parent/guardian generally initiates the assessment. In most cases, the school initiates school-based assessments. Assessments administered universally may be mandated by state or district regulations. However, more specific or individualized assessments for specialized interventions or special education programs are generally initiated by a school administrator, teacher, or school psychologist. As a result, school-based assessment may be met with additional resistance from parents or students who do not fully understand the nature and purpose of the assessment and/or who mistrust those administering the assessment. It has been found in African American students, that high cultural mistrust can negatively impact educational expectations and values (Irving and Hudley 2005). These lower expectations and educational value may have a deleterious impact on student achievement and performance on school-based assessments. Additionally, in a study of post-segregation parental engagement in schools, parents reported feelings of having to "watch over" child and teacher interactions which was attributed to a lack of trust in the teachers (Fields-Smith 2005, p. 133). With this in mind, the purpose and function of school-based assessments should be transparent and focused to promote parental support and optimal student performance.

Fortunately, CBM, as previously discussed, is an avenue for providing transparent, focused, authentic assessment of student performance. However, the influence of bias on CBM results can be a limitation to its use with African American students. This bias highlights that there is an equally, if not more, pressing need for a movement toward the implementation of "culturally-responsive" curriculum and instruction. Culturally-responsive curricula and interventions acknowledge and value the influences of culture, language, and heritage on student learning and performance and use them as tools to create meaningful educational experiences (Klinger 2005). Identifying interventions and curriculum that are culturally-responsive and have been deemed sound and empirically validated for working specifically with African American populations is a challenge. The availability of research assessing the effectiveness of interventions and curriculum specifically with African American students is limited. As a result, the outcomes of school-based assessments may not appropriately reflect student capacity if the curriculum or intervention is not culturally-responsive and has not been shown to be effective with an African American population.

In addition to being culturally responsive in curriculum and intervention development, cultural competence or responsiveness must be possessed by the educators and evaluators interpreting the school-based assessment data. It is important to note that being culturally-responsive does not entail and goes far beyond matching the racial characteristics of educators or evaluators and students. Educators who approach instruction from a culturally-responsive lens, are defined by Villegas and Lucas (2002) as "*(a) socioculturally conscious, (b) have affirming views of students from diverse backgrounds, (c) see themselves as responsible for and capable of bringing about change to make schools more equitable, (d) understand how learners construct knowledge and are capable of promoting knowledge construction, (e) know about the lives of their students, and (f) design instruction that builds on what their students already know while stretching them beyond the familiar*." With this in mind, teacher, administrator, and school psychology-training programs generally offer little explicit instruction on reading and interpreting student-based data. Furthermore, training appears to be almost nonexistent on how to interpret information gathered from school-wide assessment with diverse populations, specifically African Americans. Many teacher training programs, school psychology programs, and school administration programs address working with "diverse" populations through no more than one lecture-based course aimed to provide skills for working with all diverse populations with a large amount of time focused on working with speakers of other languages. Rarely are these programs providing insight specifically into working with African American students. It can be assumed that this method will likely fall short of addressing the social realities of working with and assessing African American students from a complete and critical lens (Cochran-Smith 2004; Loe 2000).

Conclusion

School-based assessment data is used to drive educational decisions including teacher performance evaluations, intervention effectiveness, and student placement in special education programs. School-based assessment is a valuable tool for educators, administrators, and school psychologists working with all children. Yet, when used appropriately, this method of assessment can be particularly beneficial for use with African American students. Population-based assessment supports the use of a tiered approach to designing and implementing interventions, providing early intervention services, and concentrating resources on the students who are most in need. Additionally, this approach promotes the measurement of student progress based on local or individual standards in lieu of comparison with a larger diverse group. Using school-based assessment, academic and instructional strengths and weaknesses can be identified and appropriate interventions can be implemented to promote student success.

Thorough interpretation of these assessments is paramount as the outcomes guide educational trajectories. When working with African American populations, there are additional elements to be considered, such as the cultural-responsiveness and appropriateness of the educator, curriculum, and assessment. Preferably, school personnel developing and implementing school-based assessments will possess basic skills in cultural competence and continually work to develop their capacity for considering the influence of culture, heritage, and experiences on student performance. However, challenges related to availability of assessment measures and curriculum, lack of consideration of social influence on interpretation of results, and limited training of school personnel is disheartening. Despite an increasing use of school-based assessment data to make educational decisions, research on the effectiveness of specific school-based assessments on improving educational outcomes for African American students is nascent and considerable progress needs to be made. As such there are limited school-based assessment instruments that have been found to have explicit empirical support for an exclusively African American population. Therefore, it is recommended that instruments and practices found to be sound with a more global culturally and linguistically diverse populations are considered when working with African American students.

References

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J., & Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers* (Vol. 278). Washington, DC: Economic Policy Institute.

Betts, J., Reschley, A., Pickart, M., Heistad, D., Sheran, C., & Marston, D. (2008). An examination of predictive bias for second grade reading outcomes from measures of early literacy skills in kindergarten with respect to English-language learners and ethnic subgroups. *School Psychology Quarterly, 23,* 553–570. doi:10.1037/1045–3830.23.4.553

Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps?. *The Elementary School Journal, 106*(1), 3–20.

Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The strengths and difficulties questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(6), 557–564.

Chin, J., Dowdy, E., & Quirk, M. (2013). Universal screening in middle school: Examining the behavioral and emotional screening system. *Journal of Psychoeducational Assessment, 31,* 53–60. doi:10.1177/0734282912448137.

Christ, T. J., Scullin, S., Tolbize, A., Jiban, C. L. (2008). Implications of recent research: Curriculum-based measurement of math computation. *Assessment for Effective Intervention, 33,* 198–205. doi:10.1177/1534508407313480.

Cochran-Smith, M. (2004). Taking stock in 2004 teacher education in dangerous times. *Journal of Teacher Education, 55*(1), 3–7.

Connell, J. P., Spencer, M. B., & Aber, J. L. (1994). Educational risk and resilience in African-American youth: Context, self, action, and outcomes in school. *Child development, 65*(2), 493–506.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*(3), 184–192.

Dever, B. V., Raines, T. & Dowdy, E. (2013). Referral as an entry point for disproportionality in special education: Comparing a universal screening approach to current practice. Manuscript under review.

DiStefano, C., & Morgan, G. B. (2010). Evaluation of the BESS TRS-CA using the Rasch Rating Scale model. *School Psychology Quarterly, 25*(4), 202–212. doi:10.1037/a0021509.

Doll, B., & Cummings, J. (2008). Why population-based services are essential for school mental health, and how to make them happen in your school. In B. Doll & J. Cummings (Eds.), *Transforming school mental health services: Population-based approaches to promoting the competency and wellness of children*. Thousand Oaks: Corwin.

Donovan, S., & Cross, C. T. (2002). *Minority students in special and gifted education*. Washington, DC: National Academies Press.

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health, 2,* 166–176. doi:10.1007/s12310-010-9036-3.

Dowdy, E., Dever, B., DiStefano, C., & Chin, J. (2011a). Screening for emotional and behavioral risk among students with limited english proficiency. *School Psychology Quarterly, 26*(1), 14–26. doi:10.01097/a0022072.

Dowdy, E., Doane, K., Eklund, K., & Dever, B. (2011b). A comparison of teacher nomination and screening to identify behavioral and emotional risk. *Journal of Emotional and Behavioral Disorders*. doi:10.1177/1063426611417627.

Dowdy, E., Kamphaus, R., Twyford, J., & Dever, B. (in press). Culturally competent emotional and behavioral screening. In M. Weist, N. Lever, C. Bradshaw, & J. Owens (Eds.), *Handbook of school mental health*, 2nd edn. New York: Springer.

Durlak, J. A. (1997). *Successful prevention programs for children and adolescents*. New York: Plenum.

Eklund, K., Renshaw, T. L., Dowdy, E., Jimerson, S. R., Hart, S. R., Jones, C. N., & Earhart, J. (2009). Early identification of behavioral and emotional problems in youth: Universal screening versus teacher-referral identification. *California School Psychologist, 14,* 89–95.

Evans-Hampton, T. N., Skinner, C. H., Henington, C., Sims, S., & McDaniel, C. E. (2002). An investigation of situational bias: Conspicuous and covert timing during curriculum-based measurement of mathematics across African American and Caucasian students. *School Psychology Review, 31*(4), 529–539.

Fennimore, B. S. (1997). Moving the mountain: Assessment and advocacy for children. In A. L. Goodwin (Ed.), *Assessment for equity and inclusion: Embracing all our children* (pp. 241–260). New York: Routledge.

Ferguson, C., Fuchs, D., Fuchs, L. S., & Hamlett, C. L. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58*(5), 436–450. (Retrieved from http://go.galegroup.com/ps/i.do?id=GALE%7CA11957844&v=2.1&u=callutheran&it=r&p=AONE&sw=w).

Ferri, B., & Connor, D. (2005). Tools of exclusion: Race, disability, and (re) segregated education. *The Teachers College Record, 107*(3), 453–474.

Fewster, S., & Macmillan, R. D. (2002). School-based evidence for the validity of curriculum-based measurement of reading and writing. *Remedial and Special Education, 23,* 149–156.

Fields-Smith, C. (2005). African American parents before and after Brown. *Journal of Curriculum and Supervision, 20*(2), 129–135.

Ford, D. Y., & Harmon, D. A. (2001). Equity and excellence: Providing access to gifted education for culturally diverse students. *Journal of Secondary Gifted Education, 12*(3), 141–147.

Fore, C. III, Burke, M. D., & Martin, C. (2006). Curriculum-based measurement: An emerging alternative to traditional assessment for African American children and youth. *Journal of Negro Education, 75*(1), 16–24.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57,* 488–500.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it?. *Reading Research Quarterly, 41*(1), 93–99.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9,* 20–28.

Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*(4), 507–524.

Gerber, M. M., & Semmel, M. I. (1984). Teacher as imperfect test: Reconceptualizing the referral process. *Educational Psychologist, 16,* 137–148. doi:10.1177/0022219409335217.

Glover, T., & Albers, C. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*(2), 117–135.

Gravois, T. A., & Rosenfield, S. A. (2006). Impact of instructional consultation teams on the disproportionate referral and placement of minority students in special education. *Remedial and Special Education, 27*(1), 42–52. doi:10.1177/07419325060270010501.

Green, T. D., Mcintosh, A. S., Cook-Morales, V. J., & Robinson-Zindart, C. (2005). From old schools to tomorrow's school: Psychoeducational assessment of African American students. *Remedial and Special Education, 26,* 82–92. doi:10.1177/07419325050260020301.

Harris, J. J., Brown, E. L., Ford, D. Y., & Richardson, J. W. (2004). African Americans and multicultural education: A proposed remedy for disproportionate special education placement and underinclusion in gifted education. *Education and Urban Society, 36*(3), 304–341.

Harrison, P. L., Cummings, J. A., Dawson, M., Short, R. J., Gorin, S., & Palomares, R. (2004). Responding to the needs of children, families, and schools: the 2002 multisite conference on the future of school psychology. *School Psychology Review, 33,* 12–33.

Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59,* 636–644. doi:10.1598/RT.59.7.3.

Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics, 123*(4), 1373–1414.

Hilliard III, A. G. (1991). Why we must pluralize the curriculum. *Educational Leadership, 49*(4), 12–16.

Hintze, J. M., Callahan, J. E. III, Matthews, W. J., Williams, S. A. S., & Tobin, K. G. (2002). Oral reading fluency and prediction of reading comprehension in African American and Caucasian elementary school children. *School Psychology Review, 31,* 540–553.

Hosp, M. K., & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure, 48*(1), 10–17.

Hosp, J. L., & Reschley, D. J. (2004). Disproportionate representation of minority students in special education: Academic, demographic, and economic predictors. *Exceptional Children, 70*(2), 185–199.

Irving, M. A., & Hudley, C. (2005). Cultural mistrust, academic outcome expectations, and outcome values among African American adolescent men. *Urban Education, 40*(5), 476–496.

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (Eds.). (2007). *Handbook of response to intervention: The science and practice of assessment and intervention*. New York: Springer.

Jones, J. M. (2009). Counseling with multicultural intentionality: the process of counseling and integrating client cultural variables. In J. Jones (Ed). *The psychology of multiculturalism in schools*. Bethesda: National Association of School Psychologists.

Kamphaus, R. W. (2012). Screening for behavioral and emotional risk: constructs and practicalities. *School Psychology Forum, 6,* 89–97.

Kamphaus, R. W., & Reynolds, C. R. (2007). *Behavior Assessment System for Children—Second Edition (BASC-2): Behavioral and Emotional Screening System (BESS)*. Bloomington: Pearson.

Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review, 37,* 374–390.

Kluger, R. (2011). *Simple justice: The history of Brown v. Board of Education and Black America's struggle for equality*. New York. Vintage.

Kranzler, J. H., Miller, M. D., & Jordan, L. (1999). An examination of racial/ethnic and gender bias on curriculum-based measurement of reading. *School Psychology Quarterly, 14*(3), 327–342.

Lau, M. Y., & Blatchley, L. A. (2009). A comprehensive, multidimensional approach to assessment of culturally and linguistically diverse students. In J. Jones (Ed.), *The psychology of multiculturalism in schools*. Bethesda: National Association of School Psychologists.

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45,* 163–191.

Loe, S. A. (2000). *An examination of family oriented practice and cultural diversity in school psychology: A national survey of school psychology practitioners*. ProQuest (UMI Dissertations Publishing).

Marston, D. B., Mirkin, P. K., & Deno, S. L. (1984). Curriculum-based measurement: An alternative to traditional screening, referral, and identification. *The Journal of Special Education, 18,* 109–117.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00–4769). Washington, DC: U.S. Government Printing Office.

Ortiz, S. O. (2008). Best practices in nondiscriminatory assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V*. Bethesda: National Association of School Psychologists.

Palacios, E. D., & Trivedi, P. (2009). Increasing cultural literacy: historical perspectives and cultural characteristics of minority groups. In J. Jones (Ed)., *The psychology of multiculturalism in schools*. Bethesda: National Association of School Psychologists.

Raines, T. C., Dever, B. V., Kamphaus, R. W., & Roach, A. T. (2012). Universal screening for behavioral and emotional risk: A promising method for reducing disproportionate placement in special education. *The Journal of Negro Education, 81*(3), 283–296.

Raines, T. C., Dever, B. V., Kamphaus, R. W., & Roach, A. T. (2014). Measurement equivalence of basc-2 behavior and emotional screening system student form across racial groups. Manuscript under review.

Ravens-Sieberer, U., Erhart, M., Gosch, A. and Wille, N. (2008). Mental health of children and adolescents in 12 European countries—results from the European KIDSCREEN study. *Clinical Psychological Psychotherapy, 15,* 154–163. doi:10.1002/cpp.574.

Reddy, L. A., Newman, E., De Thomas, C. A., & Chun, V. (2009). Effectiveness of school-based prevention and intervention programs for children and adolescents with emotional disturbance: A meta-analysis. *Journal of School Psychology, 47*(2), 77–99.

Reschly, D. J. (2008). School psychology paradigm shift and beyond. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V*. Bethesda: National Association of School Psychologists.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children—Second Edition (BASC-2)*. Circle Pines: AGS.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology, 45,* 193–223. doi:10.1016/j.jsp.2006.11.003.

Shinn, M. R. (1989). Identifying and defining academic problems: CBM screening and eligibility procedures. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 90–129). New York: Guilford.

Shinn, M. R. (2007). Identifying students at risk, monitoring performance, and determining eligibility within response to intervention: Research on educational need and benefit from academic intervention. *School Psychology Review, 36,* 601–617.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review, 34,* 317–342.

Skiba, R. J., Poloni-Staudinger, L., Gallini, S., Simmons, A. B., & Feggins-Azziz, R. (2006). Disparate access: The disproportionality of African American students with disabilities across educational environments. *Exceptional Children, 72,* 411–424.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: review of research. *Psychology in the Schools, 42*(8)*,* 795–819.

Terras, M. M., Thompson L. C. & Minnis, H. (2009). Dyslexia and psycho-social functioning: An exploration study of the role of self-esteem and understanding. *Dyslexia: An International Journal of Research and Practice, 15,* 304–327. doi:10.1002/dys.386

Thurber, R. S., Shinn, M. R., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review, 31,* 498–513.

Thurlow, M. L., Sinclair, M. F. & Johnson, D. R. (2002, June). *Students with disabilities who drop out of school: Implications for policy and practice* (Issue brief No. 1(2). National Center on Secondary Education Transitions.

Tilly III, W. D. (2008). The evolution of school psychology to science-based practice: Problem solving and the three-tiered model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V*. Bethesda: National Association of School Psychologists.

Tyson, E. H. (2004). Ethnic difference in using behavior rating scales to assess the mental health of children: A conceptual and psychometric critique. *Child Psychiatry and Human Development, 34,* 167–201. doi:10.1023/B:CHUD.0000014996.26276.a5.

Vannest, K. J. (2012). Implementing interventions and progress monitoring subsequent to universal screening. *School Psychology Forum, 6,* 119–136.

Van Der Heyden, A. M., & Burns, M. K. (2005). Using curriculum-based assessment and curriculum-based measurement to guide elementary mathematics instruction: Effect on individual and group accountability scores. *Assessment for Effective Intervention, 30*(3), 15–31.

Waitoller, F., Artiles, A., Cheney, D., Artiles, A., & Cheney, D. (2010). The miner’s canary: A review of overrepresentation research and explanations. *Journal of Special Education, 44*(1), 29–49. doi:10.1177/0022466908329226.

Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41*(2), 85–120.

Ysseldyke, J. E., Algozzine, R., & Thurlow, M. L. (2000). *Critical issues in special education*. Boston: Houghton Mifflin.

Zirkel, S. (2005). Ongoing issues of racial and ethnic stigma in education 50 years after Brown v. Board. *The Urban Review, 37*(2), 107–126.

Neuropsychological Assessment with African American Clients

19

Brian D. Leany, Lorraine T. Benuto and Lindsay F. Morra

Neuropsychological Assessment with African American Clients

For any neuropsychological assessment the goal is to engage the client in behavioral tasks in order to identify neurocognitive impairment and sparing. These tasks are designed to assess both specific brain cortices as well as more diffused and associated cortices, with the overarching goal of providing an assessment of insult and sparing to be used in prognoses and treatment recommendations. A consensus for the domains of neurocognitive functioning (e.g., verbal, nonverbal, executive functioning, motor functioning, etc.) as well as a general correspondence between behavior and associated cortices (for example language processing generally occurring in the left hemisphere, visuospatial processing in the right, etc.) exists across racial and ethnic groups. What is of interest for the field of psychology in general is, reports of discrepant performance reported as it relates to the pattern and level of performance among various racial and ethnic groups. Specific to this edited book, there has been much discussion related to the disparity of observations for neurocognitive performance of African Americans on a variety of neuropsychological assessment measures primarily as compared to Caucasians, but also to include other racial and ethnic groups. Thus, the goal of this chapter is to elucidate the literature related to these findings, in an effort to clarify the source of these disparities, as well as to provide empirically supported recommendations for the clinician conducting a neuropsychological assessment with an African American Client.

General Considerations for the Neuropsychological Assessment of the African American Client

There has been much discussion about test disparity and bias, and specifically as it relates to test bias for the assessment of African Americans in the USA. While there are findings to suggest differences the in performance for African Americans as compared to their Caucasian counterparts, these differences are rarely as simple as an effect of race on performance. Rather, as

---

B. D. Leany (✉)
Lake's Crossing Center, Maximum Security Facility
for Forensic Mental Health Services, 500 Galletti Way,
Sparks, NV 89504, USA
e-mail: leanyb@gmail.com

B. D. Leany · L. T. Benuto
Northcentral University, Core Faculty in the Graduate
School, 10000 E University Drive, Prescott Valley,
AZ 86314, USA

L. T. Benuto
Administrative Faculty, Department of Psychology
MS0296, University of Nevada, Reno, NV 89557, USA
e-mail: dr.benuto@gmail.com

L. F. Morra
Department of Psychology, State University of New
York, P.O. Box 6000, Binghamton, NY 13902, USA
e-mail: lmorra1@binghamton.edu

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_19,

discussed throughout this chapter, disparities of performance are best accounted for by additional demographic variables of age, education and in some instances, gender. Heaton et al. (2004, 2009) is often cited as understanding the importance for the influence of these variables in neuropsychological assessment, resulting in the creation of the Heaton norms for the Halstead Reitan Neuropsychological Test Battery, but the expansion in the diverse array of neuropsychological tests since the development of these norms warrants a more current examination of how these variables impact interpretation for African American clients. Thus, in this chapter we discuss measures common to the assessment of the traditional neuropsychological domains with African American clients, provide relevant research as it relates to performance for African American clients and make recommendations for implications in clinical practice.

Premorbid Intellectual Functioning

The capability to assess change in cognition over time or after the development of neurological injury or disease is crucial as it serves as an intra-individual yardstick that should help reduce the previously discussed intragroup differences that is more easily influenced by outside factors such as ethnicity, language and race. However, most presenting clients will not have a premorbid set of baseline data readily available, and it is thus often necessary to estimate premorbid functioning within a single assessment. To this, end various instruments and methods have been developed to provide such an estimate of 'premorbid' functioning. This includes the utilization of several demographic variables and administration of 'hold' tests that are thought to measure cognitive abilities less likely to decline with time, disease, or injury. Among commonly used hold tests, various tests are of reading ability including the National Adult Reading Test (NART), Wechsler Test of Adult Reading (WTAR), and the Wide Range Achievement Test (WRAT) reading subtest (Lezak et al. 2012; Duff 2010).

Several factors have been shown to contribute to observed differences in neuropsychological performance between African Americans and Caucasians, including reading ability (Manly et al. 2005). This may be, at least in part, because self-reported level of education has been shown to significantly overestimate reading ability in many African Americans (Baker et al. 1996; Manly et al. 2002). Therefore, research has shown that reading ability has a significant effect beyond that of education on neuropsychological performance in various domains (e.g., semantic fluency, executive functioning, memory) (Dotson et al. 2008; Manly et al. 2002). This is particularly relevant in older adults or individuals of low socioeconomic status (Dotson et al. 2009; Dotson et al. 2008), and therefore important to consider when using tests of reading ability to estimate premorbid functioning.

WRAT includes a measure of oral word reading and, assesses an individual's reading ability by asking them to read a series of words that vary in their frequency of appearance in the English language. This reading subtest has been used to estimate premorbid ability given that the performance remains relatively stable even after neurological injury or disease. While this is primarily demonstrated by comparing groups of neurologically healthy individuals and impaired groups, WRAT-3 reading performance has also shown stability within an impaired sample (Johnstone and Wilhelm 1995).

Some studies suggest that matching groups on WRAT reading performance may be more effective than matching groups on education (Manley et al. 2002; Manly et al. 2005). In fact, O'Bryant and colleagues have shown that this is particularly important in minority samples. Consistent with previously discussed findings for self reported levels of education (Baker et al. 1996; Manly et al. 2002), they found greater discrepancies between self-reported education attainment and estimated reading ability as measured by the WRAT reading subtest. Within their sample, the median reading level in African Americans was 6 years less than the mean level of education. However, in the Caucasian sample, the median reading level was only one year below mean educational attainment (O'Bryant et al. 2005). Further, reading ability, as measured by the WRAT-R reading subtest, was significantly related to the performance on measures of executive function, language, attention, and inhibition in a sample of older African Americans, while no significant relation was found between education and neuropsychological performance. Effects ranged from medium to large when comparing individuals who were high and low on reading scores (Schneider and Lichtenberg 2011).

Wechsler Test of Adult Reading

The Wechsler Test of Adult Reading (WTAR; Wechsler 2001) is an instrument that includes 50 irregularly spelled words that the individual is asked to read aloud. The words continue to get progressively more difficult. It is often used as a measure of reading ability and premorbid intelligence. The WTAR has demonstrated good internal consistency and test-retest reliability (Wechsler 2001). Scores were seen to be stable in the populations following neurological damage at two-months and five months post injury (Lezak et al. 2012).

Performance on the WTAR can be impacted by several demographic factors, including ethnicity. In a sample of individuals ranging from neurologically healthy to demented, African Americans performed significantly worse on the WTAR than Caucasians (Wagner et al. 2007). Similarly, in a study by Chin and colleagues, a sample of 244 patients with probable Alzheimer's disease received the WTAR. African Americans and Caucasians within the sample differed significantly on WTAR performance, scoring 84.2 and 105.0, respectively, in addition to the performance on various other neuropsychological measures. However, differences in neuropsychological test scores between African Americans and Caucasians no longer existed after controlling for WTAR scores, while differences remained significant when only demographic variables (including age, sex, and education) were controlled for (Chin et al. 2012). Silverberg and colleagues also demonstrated that the differences in premorbid functioning, as measured by the WTAR, account for a greater proportion of variance in neuropsychological test performance between African American and Caucasian individuals than education (Silverberg et al. 2013).

Conclusions and Recommendations for Assessing Premorbid Intellectual Functioning

The assessment of premorbid intellectual functioning via the stable cognitive abilities of reading is a valuable tool. However, as discussed above, these tests show discrepant performance between African Americans and their Caucasian counterparts. While using education-matched norms may be a viable option (in particular for the WRAT; Manley et al. 2002; Manly et al. 2005), the clinician must be aware of disparities that can yield fully one standard deviation difference in performance (e.g., the WTAR; Chin et al. 2012), but even with this difference, the WTAR still accounts for differences in performance more so than age, gender and even education (Silverberg et al. 2013). Because the source of this discrepancy between reported education and reading level is unknown, it is important that the clinician exhaust collateral information to support premorbid educational attainment. Further, it is suggested that the clinician who finds consistent discrepant information for premorbid education level, may wish to utilize an intraindividual rather than an inter-individual method of interpretation.

Attention/Vigilance, Working Memory, and Processing Speed

Although considered as separate domains of functioning, attention, working memory, and processing speed are often interrelated and sometimes hard to measure independently. Further, these abilities are often preconditions for many other cognitive abilities, and therefore are typically even further intertwined with broader constructs and assessed in measures used to measure broader constructs of cognitive abilities (e.g., memory and executive functions). Therefore, it is important to note that the variables that can influence attention, working memory, and processing speed may inadvertently have an effect on additional measures (Lezak et al. 2012).

Digit span tests are among the most commonly used instruments to assess immediate auditory memory. In Wechsler tests, Digit Span includes tasks where the individual is presented with a series of numbers that progress in length. They are either asked to repeat the sequence verbatim or to manipulate the string of numbers and put them into another sequence (e.g., reverse order of how they were initially presented). Digits forward is often considered a measure of simple attention while digits backwards requires working memory (Lezak et al. 2012). Imaging studies have shown that right dorsolateral prefrontal cortex activity is implicated in both digits forward and backwards among healthy individuals (Gerton et al. 2004).

The Wechsler Adult Intelligence Scale-III (WAIS-III) was found to have excellent psychometric data among the standardization sample. Reliability coefficients have been shown to hold up in clinical samples as well. For example, reliability coefficients were 0.92 and 0.85 for Caucasians and African Americans, respectively, among a substance use disorders sample (Ryan et al. 2000a, b). However, African Americans have been shown to perform significantly worse on Digit Span relative to Caucasians with African Americans and Caucasians scoring, on average, a scaled score of 8.2 ($SD=0.4$) and 9.7 ($SD=0.3$), respectively (Boone et al. 2007).

The Letter-Number Sequencing is a commonly used measure of working memory that requires an individual to listen to a series of randomized numbers and letters and repeat the numbers and letters in numerical and alphabetical order. Progressively, longer sequences are presented until the individual gets all three items within a trial incorrect. This is thought to be more complex than simple digit span tests (e.g., Digits forward). Therefore, Letter–Number Sequencing aims to be more sensitive to detecting attentional difficulties than other span tests (Lezak et al. 2012).

While performance on the Letter–Number Sequencing subtest of the WAIS is influenced by age and education (although this relation was non-significant after accounting for reading ability), little research has been done on the impact of race or ethnicity on performance. The implications of this relation may be significant, particularly for African American samples (Johnson et al. 2006). Although the impact of race or ethnicity on performance has not been thoroughly investigated, African Americans have been shown to have lower reading levels in several studies (Albert and Teresi 1999; Manly et al. 2002). Therefore, given the impact of reading ability on Letter-Number Sequencing performance, future investigation is warranted as African Americans could demonstrate impaired performance in working memory, which is instead a deficit in crystalized knowledge.

The Paced Auditory Serial Addition Test (PASAT) is a measure of information processing capacity. It involves the presentation of 60 pairs of numbers, and the individual is asked to add each number to the number preceding it. The PASAT includes four separate trials where the numbers are presented at different rates (ranging from one word every 0.4 s to one word every 2.4 s; Lezak et al. 2012). PASAT scores were shown to be extremely sensitive to information processing speed performance (Grownall and Wrightson 1981), and scores tend to decline with age (Spikman et al. 2000), particularly after the age of 50 (Roman et al. 1991).

Normative data for the PASAT has been provided in various samples. (Mitrushina et al. 2005; Diehr et al. 1998). Gonzalez and colleagues obtained data from the previous study samples, including the HIV Neurobehavioral Research Center, the African American Norms Project, and the Alcohol Abuse and Neuropsychological Impairment Project. This data was then used to construct a normative data set that provides demographically adjusted data based on a variety of demographic variables, including age, education, sex, and race, all of which were related to performance on the PASAT (Gonzalez et al. 2013). Demographically adjusted normative data is essential, given that research suggests that ethnicity is among significant demographic predictors of performance on the PASAT (Diehr et al. 1998).

Continuous performance tests, such as the Continuous Performance Test II, are the tests that measure vigilance or sustained attention. In order to successfully complete this task, individuals must sustain attention and remain focused on a particular target, while ignoring other distractors, over a period of time. These tasks are often computerized in order to provide information about reaction time in addition to omission and commission errors. Neural activity has primarily been observed in several frontal regions while engaging in this task (Ogg et al. 2008). Little research has been done to date on the potential relation between ethnicity and performance on continuous performance tests.

Various symbol substitution tests, such as the Symbol Digit Modalities Test (SDMT), have been used in order to measure visual scanning, sustained attention, and processing speed. A benefit of the SDMT is that, it allows for differentiation between visuomotor and oral responses to help detect where the deficit occurs in the case of impaired performance. Nine symbols, which are each paired with a number, are presented in a key. The individual must then fill in a series of empty boxes below the key that contain the symbol but area missing the associated number as quickly as they can (Lezak et al. 2012).

Research has suggested that there are differences in performance on the Symbol Digit between African Americans and Caucasians. For example, in a study of individuals aged 50–70 years old, African Americans performed significantly worse on the Symbol Digit test compared to Caucasians. However, differences between groups decreased after adjusting for several demographic variables such as literacy, occupational status, and various health characteristics (Schwartz et al. 2004). Similarly, within a sample of African Americans with a history of a traumatic brain injury, acculturation was inversely related to performance on the Symbol Digit Modalities Test, among other neuropsychological measures (Kennepohl et al. 2004). Additionally, normative data has been provided for the Modified-Symbol Digit Modalities Test for African Americans, Caribbean Black Americans, and non-Latino Whites that revealed that African Americans and non-Latino Whites performed significantly different from Caribbean Black Americans (Gonzalez et al. 2007).

An attentional task that incorporates an element of visuomotor tracking is the Trail Making Test Part A (Trails A), where the individual is asked to connect a series of circled numbers in numerical order as quickly as they can. Although administration instructions may vary slightly depending on the version administered, scores are determined by the time required to successfully connect the numbers from start to finish. This task has been shown to correlate with other visual search tasks such as Digit Symbol (Lezak et al. 2012).

Differences between African Americans and Caucasians on the Trail Making Test Part A have also been observed (Schwartz et al. 2004; Lucas et al. 2005). Schwartz and colleagues found that African American adults performed worse on this test relative to Caucasian adults (Schwartz et al. 2004). Similar results were found in a sample of patients receiving outpatient neuropsychological evaluation. On average, the African American group completed Trails A in 51.1 s while the Caucasian group completed the test in 38.6 s, after controlling for age and education (Boone et al. 2007). Mayo's Older African Americans Normative Studies provides age-corrected normative data using a sample of African Americans aged 55 and older (Lucas et al. 2005).

Recommendations for Attention/Vigilance, Working Memory, and Processing Speed

Because there are differences noted for digit span and letter number sequencing tasks, it is suggested that the tests with demographically adjusted norms such as the PASAT be incorporated into a neuropsychological assessment battery, in particular when deficits in this domain are noted. It is important to note the limited research for tasks of continuous performance as deficits in vigilance is often an indicator of neurological insult. Thus, while there is no data to suggest disparity of performance for African Americans as compared to other racial and ethnic groups, it is suggested that these measures be used with caution, and are supported with collateral information to assess intraindividual changes in performance. Finally, although differences in performance have been found for African Americans on Trails tasks (Boone et al. 2007), there are age-corrected norms available specific to African Americans (Lucas et al. 2005) that should be employed when using this measure.

Visuospatial Processing

Visuospatial processing involves a variety of cognitive abilities, including perception, organization, and visuomotor coordination, among others, and thus can be impacted by a variety of brain insults (Lezak et al. 2012). These tasks typically involve some sort of demand performance characteristic (e.g., block design or comparison of visual information) or they involve a component of learning and retrieval of visual information. We begin here by examining performance-based tasks or visuospatial processing.

Visuospatial processing has been the area where researchers have identified worse performance for African Americans relative to Caucasians, even after matching groups on years of education (Manly et al. 1998). As previously mentioned in the outset of this chapter, these discrepancies may be, in part, attributable to other biological predisposing factors (e.g., health or cardiovascular risk factors) that differ among ethnic groups. For example, visuospatial impairments were associated with cardiovascular risk in an elderly African American population after controlling for both age and education (Pugh et al. 2003). Thus, these findings for visuospatial tasks further support the need and potential importance of developing normative data based on demographic variables other than age or education in some neuropsychological measures.

As an example of the diffuse regions involved in visuospatial skills, the block design subtest of the Wechsler Adult Intelligence Scale is a constructional test requiring the manipulation of three-dimensional blocks to match a target design initially presented with target blocks and faded to stimulus cards. This task is thought to be a measure of visuospatial organization, with various other cognitive processes contributing to performance. Thus impairment on this task is often observed subsequent to neurological insult (Lezak et al. 2012). Of note, a variety of sociodemographic factors have been shown to impact block design performance.

The most notable factor identified as influencing performance on this task is age (Ryan et al. 2000a, b). However, several studies have also shown that the race has a significant effect on block design performance such that African Americans perform, on average, one point lower than Caucasians across all age levels (Kaufman et al. 1991; Marcopulos et al. 1997). However, other studies have shown that these group differences are likely attributable to other factors such as education and acculturation (Ardila and Moreno 2001; Manly et al. 1998). For example, lower levels of acculturation, as measured by the African American Acculturation Scale-Short Form inventory, were related to poorer performance on the block design in a sample of African American patients with a history of traumatic brain injury (Kennepohl et al. 2004).

An additional example of a visuoconstruction task (albeit using a different modality), the Rey-Osterrieth Complex Figure Test (ROCFT) copy trial is a test of visuoconstruction ability where the individual is asked to copy a complex figure onto a sheet of paper, with scores emphasizing both accuracy of the drawing and correct placement of individual items within the figure. The copy trial of the ROCFT has been shown to be relatively stable between the ages of 20 and 50 years old, but then begins to gradually decline (Mitrushina 2005). Thus, changes in this ability between these ages would be indicative of an organic neurocognitive problem. However, when considering administration, among healthy individuals, age, sex, and IQ are all significantly related to the ROCFT copy trial (Gallagher and Burke 2007).

Specific to the applications for African American clients, lower scores on the ROCFT have been reported in African Americans compared to Caucasians (Boone et al. 2007; Schwartz et al. 2004). Boone and colleagues found that both Caucasians ($M=31.7$, $SD=0.6$) and Asians

(*M* = 32.5, *SD* = 1.3) scored significantly higher on the ROCFT copy trial than African Americans (*M* = 26.9, *SD* = 1.0), and these differences in raw scores between African Americans and the other groups were clinically significant (2007). Scores for the African American group, on average, fell below the first percentile compared to normative data (Meyers and Meyers 1995). Of primary import, group differences decreased after controlling for various sociodemographic factors, including literacy, occupational status, and demographic variables (Schwartz et al. 2004). Similarly, in other studies, education level has been shown to impact performance on the ROCFT copy trial within various cultural groups including African Americans (Unverzagt et al. 1996). Again, these results suggest that socio-cultural factors that come into play are more specific to challenges for African American's ability to attain relevant skills and status rather than test bias.

The Judgment of Line Orientation (JLO) measures one's ability to estimate angular relationships between a series of lines that form a semicircle. The test includes 30 items, and the individual is asked to match a pair of lines to the two matching lines within the semicircle (numbered one to eleven) (Lezak et al. 2012). Performance in patients with left hemisphere damage tends to remain relatively intact, while patients with right hemisphere lesions are more likely to show impaired performance (Benton et al. 1994). Scores on the JLO are demographically adjusted for both age and sex, given that the performance has been shown to decline as one ages and women tend to perform, on average, approximately two points lower than men (Lezak et al. 2012).

The Mayo's Older African American Normative Studies (MOAANS) have attempted to create a database of normative data for various neuropsychological tests that are ethnicity specific. Within their studies they have created two normative data sets for the JLO in older African Americans (aged 55 and older). One data set includes data from all participants, including those who failed sample items. The second data set is collected from only a subset of individuals who passed the sample items (Lucas et al. 2005). Thus, when assessing visuospatial abilities for the African American, it is of great importance to use African American specific norms (when available, such as those for the JLO), with emphasis on accounting for age, literacy and educational attainment.

Recommendations for Visuospatial Processing

Similar to other domains, many of the tasks described above can be reliably employed in neuropsychological assessment when accounting for age, education, and literacy. Thus, the literature for this domain reinforces the importance of accurately assessing educational attainment and reading abilities, whose challenge is discussed in the outset of this chapter. It should be noted that, there is evidence that suggest additional predisposing factors which may contribute to racial and ethnic differences, specifically as it relates to visuospatial processing (e.g., the aforementioned contribution of prevalence rates for cardiovascular disease; Pugh et al. 2003), and thus, tests that lack an appropriate inclusion of African Americans in the normative sample, as well as those that lack norms that account for age and education, should be used with some trepidation, in particular for this domain.

Visual Learning and Memory

In addition to the performance based tasks of visual tasks, we must assess the individual's abilities related to visual learning and memory. Many of these tasks are analogs to the verbal learning and memory tasks and thus administration and scoring procedures should feel familiar.

As previously discussed, the ROCFT includes a measure of visual learning and memory, but there are more specific measures of visual learning and memory that do not require reproduction, but rather rely on recognition and identification, thus reducing some of the cognitive load for associative cortices required for reproduction.

As a battery, the Wechsler Memory System-IV (WMS-IV; Wechsler 2009) includes an index of visual memory, including immediate and delayed visual memory, though not a direct measure of visual learning (as with other measures discussed below). One of the goals for the development of the latest iteration of this measure was to create a more culture neutral test. Further, subsequent research comparing individuals with mild TBI, moderate TBI and controls, that included a representative sample of African Americans, found no differences among ethnic groups for performance on the WMS-IV visual memory index (Carlozzi et al. 2013).

An example of the analog between visual and auditory learning and memory tasks, the Biber Figure Learning Test-Extended (BFLTE; Glosser et al. 1997) uses a paradigm similar to that of the CVLT (Glosser et al. 2002; Kurtzman 1996; Tracy et al. 2001; discussed below), to assess learning and memory. It assesses immediate and delayed recall for novel geometric shapes, and includes interference as well as recognition tasks. Although the CVLT and the BFLT-E are not identically matched in terms of difficulty level and item content, they can serve as relative measures of verbal and non-verbal learning (Tracy et al. 2001). The psychometrics for this test as well as its ease of administration makes it attractive. Inter-rater reliability for the BLFT-E has been found to be as high as 0.98, with good test-retest reliability as well as good criterion validity (Glosser et al. 2002). There is little data that specifically examines performance for African Americans as compared to other racial or ethnic groups however, available research that include African American participants, and reports analyses for differences among ethnic groups does not indicate a difference in performance for African Americans (e.g., Duke et al. 2010; Mattson and Roebuck 2002)

The Brief Visuospatial Memory Test-Revised (BVMT-R; Cherner et al. 2007) is a multi-trial presentation of geometric figures with immediate, delayed, and recognition tasks. The authors report solid psychometrics for the measure (with reliability coefficients that range from 0.96 to 0.97 for the learning trials, 0.97 for Total Recall, and 0.97 for Delayed Recall. They further report that the Test-retest reliability coefficients range from 0.60 for Trial 1 to 0.84 for Trial 3; Cherner et al. 2007). Researchers report that, a cursory examination of differences in performance between Caucasians and African Americans on the BVMT-R would suggest effects of race, with Caucasians outperforming African Americans on learning and recall trials. However, a regression analysis indicates that the most prominent influence on performance for age, with older African Americans performing more poorly than their Caucasian counterparts (Norman et al. 2011). Most importantly, they note that the raw score differences do not appear large ($M=1.5$), but classification errors in impairment are greater, in particular when looking at scores for the delayed recall component (raw scores: Caucasians $M=10.2$, $SD=1.7$; African American $M=8.7$, $SD=2.4$). Thus Norman and colleagues (2011) suggest using their corrected norms when using this measure to Assess African Americans.

The Recognition Memory Test (RMT; Warrington 1984) is one that includes tasks that require a memory of words and a memory of faces task. Its ease of administration and low demand characteristics make it ideal for assessing individuals who may have speech production or associative language cortex deficits. It includes an exposure/learning presentation and then a test condition with the target and a distractor presented. For this test, researchers were unable to find differences in performance between Caucasians and African Americans (O'Bryant et al. 2003). This is important, especially considering that this test is reported to be specific to memory function (as opposed to attention, concentration, etc.) as well as less susceptible to secondary effects of anxiety and depression (Warrington 1984), and thus would indicate a possibly unbiased measure of pure visual memory.

Recommendations for Visual Learning and Memory

Assessing visual learning and memory appears to be less fraught with challenges than other domains. Many of the tests used have good psychometric properties that appear to be free from racial or ethnic bias. Additionally, newer measures of visual learning and memory such as those found in the latest version of the WMS (Wechsler 2009), appear to have met their goal of reducing racial and cultural bias. While existing measures, such as the BVMT-R, have been studied empirically to examine racial differences, with researchers generating normative data specific to African Americans that accounts for the effects of age (Norman et al. 2009). Thus, the clinician conducting a neuropsychological assessment with an African American client, should feel confident in their ability to identify psychometrically and clinically sound measures of visual learning and memory.

Language

Language may be among the most difficult of cognitive domains to assess cross culturally, given that the performance on language measures is largely influenced by both education and culture (Manly et al. 2004; Gao et al. 2007). Language based assessment consists of generally performance based tasks such as fluency and naming, but can also include learning and memory based tasks.

Snitz et al. (2009) investigated the effect of various socio-demographic variables, including age, gender, education, and race on language ability. Their sample consisted of individuals from two large epidemiological studies (Monongahela-Youghiogheny Healthy Aging Team and Indianapolis Study of Health and Aging). Additionally, they developed normative data for some the Animal Fluency Task and the Indiana University Token Test that included a representative group of African Americans that varied on a range of socio-demographic variables, making this a good source of normative data (Snitz et al. 2009) when using these tests to assess African Americans.

The Boston Naming Test (BNT) is a measure of visual confrontation naming in which the individual is asked to look at and name up to 60 line drawings. If the individual is unable to spontaneous provide the name, they are presented with a series of cues (semantic and/or phonemic) to aid them in confrontation naming. Finally, if the individual is still unable to name the item after being given cues, they are provided with four choices (Lezak et al. 2012). Various neuroanatomical lesions are associated with dysnomia. For example, semantic errors are associated with lesions in the posterior superior temporal and inferior parietal regions. Additionally, lesions of the insula and putamen are associated with increased phonological errors (Knopman et al. 1984).

Differential performance between African Americans and Caucasians has been observed in several studies (Boone et al. 2007; Manly et al. 1998; Wagner et al. 2007). For example, Pedraza and colleagues evaluated BNT performance in individuals who participated in the Mayo's Older African Americans Normative Studies (MOAANS) and found that Caucasian adults performed, on average, significantly higher than African American adults (groups means were 52.9 and 43.3, respectively; Pedraza et al. 2009). Therefore, some have cautioned that the BNT may not be an appropriate test in ethnic minorities. While various explanations have been proposed to describe the relation between ethnicity and performance, including differences in education (Manley et al. 2002) and cultural appropriateness of the items (Whitfield et al. 2000), some studies have shown that differences in group performance still remain significant after controlling for reading ability (Baird et al. 2007). Therefore, the development of normative data corrected for ethnicity, among other demographic variables, may be useful in minimizing the impact of these sources of variance. Such normative data for the BNT has been compiled within a number of studies specific to various groups of African American individuals (Fillenbaum et al. 1997; Ross and Lichtenberg 1998; Lucas et al. 2005; Manly et al. 1998; Unverzagt et al. 1996).

In addition to naming tasks, measures of fluency are often used to assess language (e.g., Controlled Oral Word Association Test; COWAT) the COWAT measures one's ability to generate words quickly that begin with a certain letter (e.g., F, A, and S). Individuals are given

60 s for each letter and are asked to produce as many words as possible without using any proper names or providing the same word with different endings (Benton et al. 1994). Similarly, in measures of category fluency, individuals are asked to produce as many words as possible within 60 s. However, with category fluency test, they are provided a semantic category (e.g., animals, fruits, and vegetables) and generated words can begin with any letter. Impairment on measures of word fluency is often associated with left frontal lesions (reviewed in Lezack et al. 2012).

While there is a paucity of research studying the impact of racial and ethnic differences on measures of verbal fluency, some research has shown that African Americans perform worse on these tests relative to Caucasians. In a study of 600 male veterans aged 31–46, there was a significant main effect for ethnicity after controlling for education and income for both letter and animal fluency (Johnson-Selfridge et al. 1998). Additionally, older African Americans have also been shown to obtain lower scores compared to Caucasians on category fluency (Manly et al. 1998). Subsequently, some studies aimed to provide normative data for African Americans. Gladsjo and colleagues collected data from a normative sample of adults who were participating in the African American Norms Project (AANP), and found that education, age and ethnicity all accounted for a significant proportion of the variance in letter and category fluency performance (Gladsjo et al. 1999). The Mayo's Older African American Normative Studies (MOAANS) provides age-adjusted and education corrected normative data for older adults (age 55 and older) on the COWAT (Lucas et al. 2005).

Verbal Learning and Memory

Verbal learning tasks most frequently consist of the examiner reading lists of words to the client who then repeats the words back under various cues via free recall at various time intervals. Most commonly cited within the literature reviewed were the California Verbal Learning Test, the Auditory Verbal Learning Test, and the Hopkins Verbal Learning Test. Adjusted norms for these tests are based on demographic factors including age, years of education, and sex exist. In fact the development of demographically adjusted norms have been the focus of a moving body of literature that highlights the manner in which demographic factors are related to test performance

Specific to ethnicity and race, some researchers have found between-ethnic group performance on measures of verbal learning although the majority of researchers note that once other demographic factors are controlled for between-ethnic group differences disappear. For example, Manly and colleagues (1998) examined neuropsychological test performance among a randomly selected community sample of English-speaking non-Hispanic African American elders who had been determined to be cognitively healthy (i.e., no indication of dementia) by a neurologist. African American elders obtained significantly lower scores on measures of verbal and nonverbal learning and memory, abstract reasoning, language, and visuospatial skill than the Caucasians. However, once differences in education were accounted for significant ethnic group differences on measures of figure memory, verbal abstraction, category fluency, and visuospatial skill only remained. Similarly, Fillenbaum et al. (2001) investigated the CERAD protocol, which includes a list-learning task to examine the extent to which ethnic differences were present once education and age adjustments were made. While African Americans tended to perform poorer than their Caucasian counterparts, when demographic characteristics were controlled for (e.g., education, age) racial differences were not present. Indeed, Friedman et al. (2002) noted that the demographic adjustments may be more important for African Americans than the Caucasians. Specifically, Friedman and colleagues examined the influence of demographic characteristics on the Hopkins Verbal Learning Test in a sample of 237 African American elders. From this, they developed normative data stratified by age with score adjustments for education and gender. They noted that in their sample, education was related to performance but in samples of Caucasians this has not been noted to be the case (i.e., Vanderploeg et al.

2000). Beyond between-ethnic group differences, researchers have also examined within group differences among African Americans. Specifically, Kennepohl and colleagues (2004) found that the lower levels of acculturation were associated with poorer performance on the Rey Auditory Verbal Learning test among African Americans who had experienced a Traumatic Brain Injury. These findings highlight the possibility of within group differences and suggest that clinicians may need to consider the acculturation level of their client. Indeed, this is a consideration that is frequently recommended in the psychological assessment literature (Benuto 2013; Benuto et al. In Press) but most typically discussed with regard to Asian and Hispanic groups. Nonetheless, measures that assess African American acculturation do exist and Kennepohl et al.'s findings suggest that at the very least this phenomenon should be further researched.

Despite that several researchers have noted that once education and age are accounted for, racial differences in test performance are not present, norms for African Americans have been developed. Specifically, Norman and colleagues (2000) examined the California Verbal Learning Test (CVLT) performance among African American and Caucasian participants. Age, education, ethnicity, and gender were found to be significant predictors of performance on several indices on the CVLT and as such demographically corrected scores were developed. With regard to the Auditory Verbal Learning Test, Ferman et al. (2007) reported on data from the Mayo Older African Americans Normative Studies and established age and education adjusted norms for use with African American elders. More recently, Norman et al. (2011) developed Hopkins Verbal Learning Test-Revised (HVLT-R) norms for both African Americans and Caucasians noting that previously published normative standards significantly overestimated impairment in African Americans. The authors noted that when the norms they developed are used, age, education, gender are not related to performance on these measures. They also noted that, the magnitude of demographic contributions (especially age) was greater for African Americans than for Caucasians on most measures.

Recommendations for Auditory Learning and Memory Tasks

Based on the information discussed above it is apparent that when using verbal learning tests, the clinician should carefully examine the extent to which age and education (at the very least) might impact test performance and should use adjusted scores. Given that normative data is available for a number of verbal tests, the clinician is highly encouraged to utilize these adjusted norms. Specifically, the clinician is encouraged to optimally consider the Auditory Verbal Learning Test (see Ferman et al. 2007 for norms and adjustments); the HVLT-R (see Norman et al. 2011 for norms and adjustments); and the CVLT (Norman et al. 2000 for norms and adjustments) for use with the African American client when selecting a verbal learning test.

Conclusions and Recommendations

In spite of the challenges discussed above, there is a great deal of information available, specifically appropriate normative data that would allow a reliable and valid assessment for African American clients that has been summarized in this chapter (See Table 19.1 for a summary of measures and related research). Common to many measures used to assess neuropsychological functioning in African Americans is the influence of age as well as education (to some extent; literacy and gender) as variables that influence prediction and classification of impairment or sparing, when assessing African American clients. Each of the domains discussed was able to identify appropriate tests, with corresponding normative data. In some instances the research created new norms to use in assessing performance with African Americans (e.g., Norman et al. 2011). The most challenging domain, and not an unimportant one, was that of premorbid intellectual functioning and reading abilities. For this domain, it was suggested that

Table 19.1 Summary of neuropsychological measures

Domain(s) assessed	Test name	Research and recommendations
Premorbid IQ	*Wechsler test of adult reading (WTAR)*	Performance for African Americans is significantly worse than for Caucasians (Wagner et al. 2007), but appeared to account for findings of significant differences in neuropsychological test performance between the groups that was not accounted for by other demographic variables(Chin et al. 2012; Silverberg et al. 2013)
	Wide range achievement test (WRAT) oral word subtest	This measure consistently demonstrates reading levels discrepant with reported educational attainment in African Americans (Baker et al. 1996; Manly et al. 2002; O'Bryant et al. 2005), and is significantly, positively correlated with neuropsychological performance domains of executive functioning, language, attention and inhibition (Schneider and Lichtenberg 2011)
Attention/vigilance, working memory, and processing speed	*Digit span*	African Americans have been shown to perform significantly worse relative to Caucasians with African Americans and Caucasians scoring, on average, a scaled score of 8.2 ($SD=0.4$) and 9.7 ($SD=0.3$), respectively (Boone et al. 2007)
	Letter-number sequencing	Although the impact of race or ethnicity on performance has not been thoroughly investigated, African Americans have been shown to have lower reading levels in several studies (Albert and Teresi 1999; Manly et al. 2002)
	Paced auditory serial addition test (PASAT)	Normative datahas been provided in various samples (Mitrushina et al. 2005; Diehr et al. 1998) and, given that research suggests that ethnicity is among significant demographic predictors of performance (Diehr et al. 1998) are essential
	Symbol digit modalities test (SDMT)	Poorer performance has been observed for older African Americans, but with decreased differences after adjusting for literacy, occupational status, and various health characteristics (Schwartz et al. 2004). Those with a history of TBI, demonstrated an inverse relation between acculturation and perfotrmance on this task, among other neuropsychological measures (Kennepohl et al. 2004)
	Trail making test part A (Trails A)	African American adults performed worse on this test relative to Caucasian adults (Schwartz et al. 2004), which has also been found in patients receiving outpatient neuropsychological evaluation even after controlling for age and education (Boone et al. 2007). MOAANs Studies provides age-corrected normative data using a sample of African Americans aged 55 and older (Lucas et al. 2005)
Visuospatial processing	*Block design*	Though race is reported to have a significant effect on performance (Kaufman et al. 1991; Marcopulos et al. 1997), others found that differences are likely due to education and acculturation (Ardila and Moreno 2001; Manly et al. 1998), with lower levels of acculturation correlating with poorer performance in a TBI sample (Kennepohl et al. 2004), and age as the most influential predictor of performance (Ryan et al. 2000a, b)
	Rey-osterrieth complex figure test (ROCFT)	Boone and colleagues found that both Caucasians ($M=31.7$, $SD=0.6$) and Asians ($M=32.5$, $SD=1.3$) scored significantly higher on the ROCFT copy trial than African Americans ($M=26.9$, $SD=1.0$), and these differences in raw scores between African Americans and the other groups were clinically significant (2007). Differences decrease after controlling for literacy, occupational status, and demographic variables (Schwartz et al. 2004). Education level has been shown to impact performance for the copy trial for African Americans (Unverzagt et al. 1996)

Table 19.1 *(continued)*

Domain(s) assessed	Test name	Research and recommendations
	Judgment of line orientation (JLO)	MOAANS normative data, including two for the JLO in older African Americans exists. One set includes data form all participants, including those who failed sample items. The second data set contains only individuals who passed the sample items (Lucas et al. 2005). One should be cognizant of which data set is being used
Visual learning and memory	*Wechsler memory system-IV (WMS-IV), visual memory index*	Research comparing TBI and Controls, that included a representative sample of African Americans, found no differences among ethnic groups (Carlozzi et al. 2013)
	Biber figure learning test-extended (BFLTE)	Available research indicates no racial or ethnic differences (e.g., Duke et al. 2010; Mattson and Roebuck 2002)
	Brief visuospatial memory test–revised (BVMT–R)	A regression analysis indicates age rather than race accounts for poorer performance in African Americans when compared to their Caucasian counterparts (Norman, et al. 2011) as well as note that although the raw score differences do not appear large ($M=1.5$), classification errors in impairment are greater, especially for delayed recall, and suggest using their corrected norms (Norman et al.2011)
	Recognition memory test	Researchers found no differences for race (O'Bryantet al. 2003). This test is reported to be specific to memory function, and less susceptible to secondary effects of anxiety and depression (Warrington 1984)
Language	*Animal fluency task and the Indiana university token test*	Normative data for these measures that includes Afriacn Americans has been developed and should be used for these tests (Snitz et al. 2009)
	Boston naming test (BNT)	Differences in group performance persist even after controlling for reading ability (Baird et al. 2007). Therefore, normative data has been compiled within a number of studies specific to various groups of African American individuals (Fillenbaum et al. 1997; Unverzagt et al. 1996; Manly et al. 1998; Lichtenberg et al. 1998; Ross and Lichtenberg 1998; Lucas et al. 2005)
	Controlled oral word association test (COWAT)	African Americans perform worse relative to Caucasians, with a significant main effect for ethnicity even after controlling for education and income for both letter and animal fluency (Johnson-Selfridge et al. 1998), and obtain lower scores on category fluency (Manly et al. 1998). Education, age and ethnicity also have been shown to account for a significant proportion of the variance in letter and category fluency performance (Gladsjo et al. 1999). MOAANS provides age-adjusted and education corrected normative data for older adults (age 55 and older) on the COWAT (Lucas et al. 2005)
Verbal learning and memory	*California verbal learning test (CVLT)*	Researchers found that demographically adjusted norms that account for age and education significantly reduce or eliminate effects of race and ethnicity (Fillenbaum et al. 2001; Manly et al. 1998)
	Hopkins verbal learning test-revised (HVLT-R)	Friedman et al. (2002) noted demographic adjustments may be more important for African Americans than Caucasians, and developed normative data. They noted that in their sample, education was related to performance but in samples of Caucasians this has not been the case (i.e., Vanderploeg et al. 2000)

collateral information be used in conjunction with obtained scores to best assess changes in level of performance. Further, while not specifically discussed in this chapter, there are efforts to collect demographically corrected norms for specific areas of neuropsychological insult, such as Alzheimer's dementia (e.g., Mayo's Older African Americans Normative Studies; Lucas et al. 2005), which continue to refine the clinician's diagnostic and thus prognostic abilities. There is no doubt that endeavors such as this will continue, and clinicians should keep abreast of emerging findings as they relate to neuropsychological assessment of the African American client.

References

Albert, S. M., & Teresi, J. A. (1999). Reading ability, education, and cognitive status assessment among older adults in Harlem, New York City. *American Journal of Public Health, 89,* 95–97.

Ardila, A., & Moreno, S. (2001). Neuropsychological test performance in Aruaco Indians: An exploratory study. *Journal of the International Neuropsychological society, 7,* 510–515.

Baird, A. D., Ford, M., & Podell, K. (2007). Ethnic differences in functional and neuropsychological test performance in older adults. *Archives of Clinical Neuropsychology, 22*(3), 309–318.

Baker, F. M., Johnson, J. T., Velli, S. A., & Wiley, C. (1996). Congruence between education and reading levels of older persons. *Psychiatric Services, 47,* 194–196.

Benedict, R. (1997). *Brief visuospatial memory test-revised*. Odessa: Psychological Assessment Resources

Benton, A. L., Hamsher, K. D., Varney, N. R., & Spreen, O. (1983). *Judgment of line orientation*. New York: Oxford University Press.

Benuto, L. T., & Leany, B. (2013). Assessing effort & malingering with the hispanic client. In L. T. Benuto (Ed.), *Guide to psychological assessment with hispanics*. New York: Springer.

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology, 22,* 355–365.

Carlozzi, N. E., Grech, J., & Tulsky, D. S. (2013). Memory functioning in individuals with traumatic brain injury: An examination of the Wechsler Memory Scale-fourth edition (WMS-IV), *Journal of Clinical and Experimental Neuropsychology, 35*(9), 906–914, doi:10.1080/13803395.2013.833178.

Cherner, M., Suarez, P., Lazzaretto, D., Fortuny, L., Mindt, M. R., Dawes, S., & Heaton, R. (2007). Demographically corrected norms for the brief visuospatial memory test-revised and hopkins verbal learning test-revised in monolingual Spanish speakers from the US–Mexico border region. *Archives of Clinical Neuropsychology, 22*(3), 343–353.

Chin, A. L., Negash, S., Xie, S., Arnold, S. E., & Hamilton, R. (2012). Quality, and not just quantity, of education accounts for differences in psychometric performance between African Americans and white non-hispanics with Alzheimer's disease. *Journal of the International Neuropsychological Society, 18*(2), 277–285.

Diehr, M. C., Heaton, R. K., Miller, W., & Grant, I. (1998). The paced auditory serial addition task (PASAT): Norms for age, education, and ethnicity. *Assessment, 5*(4), 375–387.

Duff, K. (2010). Predicting premorbid memory functioning in older adults. *Applied Neuropsychology, 17*(4), 278–282.

Duke, L. A., Allen, D. N., Ross, S. A., Strauss, G. P., & Schwartz, J. (2010). Neurocognitive function in schizophrenia with comorbid posttraumatic stress disorder. *Journal of Clinical & Experimental Neuropsychology, 32*(7), 737–751. doi:10.1080/13803390903512660.

Fillenbaum, G. G., Heyman, A., Huber, M. S., Ganguli, M., & Unverzagt, F. W. (2001). Performance of elderly African American and white community residents on the CERAD neuropsychological battery. *Journal of the International Neuropsychological Society, 7*(4), 502–509.

Fillenbaum, G. G., Huber, M., & Taussig, I. M. (1997). Performance of elderly White and African American community residents on the abbreviated CERAD Boston naming test. *Journal of Clinical and Experimental Neuropsychology, 19*(2), 204–210.

Friedman, M. A., Schinka, J. A., Mortimer, J. A., & Graves, A. B. (2002). Hopkins verbal learning test–revised: Norms for elderly African Americans. *The Clinical Neuropsychologist, 16*(3), 356–372.

Ferman, T. J., Lucas, J. A., Ivnik, R. J., Smith, G. E., Willis, F. B., Petersen, R. C., & Graff-Radford, N. R. (2005). Mayo's older African American normative studies: Auditory verbal learning test norms for African American elders. *The Clinical Neuropsychologist, 19*(2), 214–228.

Gallagher, C., & Burke, T. (2007). Age, gender and IQ effects on the rey-osterrieth complex figure test. *British Journal of Clinical Psychology, 46*(1), 35–45

Gao, S., Jin, Y., Hall, K. S., Liang, C., Unverzagt, F. W., Ji, R., & Hendrie, H. C. (2007). Selenium level and cognitive function in rural elderly Chinese. *American Journal of Epidemiology, 165*(8), 955–965.

Gladsjo, J. A., Schuman, C. C., Evans, J. D., Peavy, G. M., Miller, S. W., & Heaton, R. K. (1999). Norms for letter and category fluency: Demographic corrections for age, education, and ethnicity. *Assessment, 6*(2), 147–178.

Glosser, G., Cole, L., Khatri, U., DellaPietra, L., & Kaplan, E. (2002). Assessing nonverbal memory with the biber figure learning test—extended in temporal lobe epilepsy patients. *Archives of Clinical Neuropsychology, 17*(1), 25–35.

Glosser, G., Cole, L. C., French, J. A., Saykin, A. J., & Sperling, M. R. (1997). Predictors of intellectual performance in adults with intractable temporal lobe epilepsy. *Journal of the International Neuropsychological Society, 3*(3), 252–259.

Gerton, B. K., Brown, T. T., Meyer-Lindenberg, A., Kohn, P., Holt, J. L., Olsen, R. K., & Berman, K. F. (2004). Shared and distinct neurophysiological components of the digits forward and backward tasks as revealed by functional neuroimaging. *Neuropsychologia, 42*(13), 1781–1787.

Gonzalez, H. M., Whitfield, K. E., West, B. T., Williams, D. R., Lichtenberg, P. A., & Jackson, J. S. (2007). Modified-symbol digit modalities test for African Americans, Caribbean Black Americans, and non-Latino Whites: Nationally representative normative data from the National Survey of American Life. *Archives of Clinical Neuropsychology, 2,* 605–613.

Gonzalez, R., Grant, I., Miller, S. W., Taylor, M. J., Schweinsburg, B. C., Carey, C. L., et al. (2013). Demographically adjusted normative standards for new indices of performance on the Paced Auditory Serial Addition Task. *The Clinical Neuropsychologist, 20*(3), 396–413.

Gronwall, D. M. A., & Wrightson, P. (1981). Memory and information processing capacity after closed head injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 44,* 889–895.

Heaton R. K, Miller S. W, Taylor M. J, & Grant I. (2004). *Revised comprehensive norms for an expanded Halstead Reitan battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults*. Lutz: Psychological Assessment Resources.

Heaton R. K., Ryan, L., & Grant, I. (2009). Demographic influences and use of demographically corrected norms in neuropsychologial assessment. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric and neuromedical disorders* (pp. 127–155). New York: Oxford University Press.

Johnson, A. S., Flicker, L. J., & Lichtenberg, P. A. (2006). Reading ability mediates the relationship between education and executive function tasks. *Journal of the International Neuropsychological Society, 12,* 64–71.

Johnson-Selfridge, M. T., Zalewski, C., & Aboudarham, J. F. (1998). The relationship between ethnicity and word fluency. *Archives of Clinical Neuropsychology, 13*(3), 319–325.

Johnstone, B., & Wilhelm, K. L. (1996). The longitudinal stability of the WRAT-R reading subtest: Is it an appropriate estimate of premorbid intelligence? *Journal of the International Neuropsychological Society, 2*(4), 282–285.

Kaufman, A. S., McLean, J. E., & Reynolds, C. R. (1991). Analysis of WAIS-R factor patterns by sex and race. *Journal of Clinical Psychology, 47,* 548–557.

Kennepohl, S., Shore, D., Nabors, N., & Hanks, R. (2004). African American acculturation and neuropsychological test performance following traumatic brain injury. *Journal of the International Neuropsychological Society, 10,* 566–577.

Knopman, D. S., Selnes, O. A., Niccum, N., & Rubens, A. B. (1984). Recovery of naming in aphasia relationship to fluency, comprehension and CT findings. *Neurology, 34*(11), 1461–1461.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York: Oxford University Press.

Lucas, J. A., Ivnik, R. J., Smith, G. E., Ferman, T. J., Willis, F. B., Petersen, R. C., et al. (2005). Mayo's older African Americans normative studies: Norms for Boston Naming Test, Controlled Oral Word Association, Category Fluency, Animal Naming, Token Test, Wrat-3 Reading, Trail Making Test, Stroop Test, and Judgment of Line Orientation. *The Clinical Neuropsychologist, 19*(2), 243–269.

Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., et al. (1998). Cognitive test performance among nondemented elderly African Americans and whites. *Neurology, 50*(5), 1238–1245.

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates difference in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society, 8,* 341–348.

Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology, 11*(1), 37–46.

Manly, J. J. (2005). Advantages and disadvantages of separate norms for African Americans. *The Clinical Neuropsychologist, 19*(2), 270–275.

Marcopulos, B. A., McLain, C. A., & Giuliano, A. J. (1997). Cognitive impairment or inadequate norms: A study of healthy, rural, older adults with limited education. *The Clinical Neuropsychologist, 11,* 111–131.

Mattson, S., & Roebuck, T. (2002). Acquisition and retention of verbal and nonverbal information in children with heavy prenatal alcohol exposure. *Alcoholism: Clinical and Experimental Research, 26*(6), 875–882.

Meyers, J. E., & Meyers, K. R. (1995). *Rey complex figure test and recognition trial manual*. United States: Psychological Assessment Resrouces.

Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment: Second Edition*. New York: Oxford University Press.

Norman, M. A., Evans, J. D., Miller, W. S., & Heaton, R. K. (2000). Demographically corrected norms for the California verbal learning test. Journal of Clinical and Experimental Neuropsychology, 22(1), 80–94.

Norman, M. A., Moore, D. J., Taylor, M., Franklin Jr., D., Cysique, L., Ake, C., & HNRC Group. (2011). Demographically corrected norms for African Americans and caucasians on the hopkins verbal learning test–revised, brief visuospatial memory test–revised, stroop color and word test, and wisconsin card sorting test 64-card version. *Journal of Clinical and Experimental neuropsychology, 33*(7), 793–804.

O'Bryant, S. E., Hilsabeck, R. C., McCaffrey, R. J., & Drew Gouvier, W. M. (2003). The recognition memory test examination of ethnic differences and norm validity. *Archives of Clinical Neuropsychology, 18*(2), 135–143.

O'Bryant, S. E., Schrimsher, G. W., & O'Jile, J. R. (2005). Discrepancies between self-reported years of education and estimated reading level: Potential implications for neuropsychologists. *Applied Neuropsychology, 12*(1), 5–11.

Ogg, R. J., Zou, P., Allen, D. N., Hutchins, S. B., Dutkiewicz, R. M., & Mulhern, R. K. (2008). Neural correlates of a clinical continuous performance test. *Magnetic Resonance Imaging, 26*(4), 504–512.

Pedraza, O., Graff-Radford, N. R., Smith, G. E., Ivnik, R. J., Willis, F. B., Petersen, R. C., & Lucas, J. A. (2009). Differential item functioning of the Boston naming test in cognitively normal African American and caucasian older adults. *Journal of the International Neuropsychological Society, 15*(5), 758–768.

Pugh, K. G., Kiely, D. K., Milberg, W. P., & Lipsitz, L. A. (2003). Selective impairment of frontal-executive cognitive function in African Americans with cardiovascular risk factors. *Journal of American Geriatric Society, 51*(10), 1439–1444.

Roman, D. D., Edwall, G. E., Buchanan, R. J., & Patton, J. H. (1991). Extended norms for the Paced Auditory Serial Addition Task. *The Clinical Neuropsychologist, 5,* 33–40.

Ross, T. P., & Lichtenberg, P. A. (1998). Expanded normative data for the Boston naming test for use with urban, elderly medical patients. *The Clinical Neuropsychologist, 12*(4), 475–481.

Ryan, J. J., Arb, J. D., Paul, C. A., & Kreiner, D. S. (2000a). Reliability of the WAIS-III subtests, indexes, and IQs in individuals with substance abuse disorders. *Assessment, 7*(20), 151–156.

Ryan, J. J., Sattler, J. M., & Lopez, S. J. (2000b). Age effects on Wechsler Adult Intelligence Scale-III subtests. *Archives of Clinical Neuropsychology, 15,* 311–317.

Schneider, B. C., & Lichtenberg, P. A. (2011). Influence of reading ability on neuropsychological performance in African American elders. *Archives of clinical neuropsychology, 26*(7), 624–631.

Schwartz, B. S., Glass, T. A., Bolla, K. I., Sewart, W. F., Glass, G., Rasmussen, M., et al. (2004). Disparities in cognitive functioning by race/ethnicity in the Baltimore Memory Study. *Environmental Health Perspectives, 112*(3), 314–320.

Silverberg, N. D., Hanks, R. A., & Tompkins, S. C. (2013). Education quality, reading recognition, and racial differences in the neuropsychological outcome from traumatic brain injury. *Archives of Clinical Neuropsychology, 28*(5), 485–491.

Snitz, B. E., Unverzagt, F. W., Chang, C. C. H., Bilt, J. V., Gao, S., Saxton, J., & Ganguli, M. (2009). Effects of age, gender, education and race on two tests of language ability in community-based older adults. *International Psychogeriatrics, 21*(6), 1051–1062.

Spikman, J. M., Deelman, B. G., & van Zomeren, A. H. (2000). Executive functioning, attention and frontal lesions in patients with chronic CHI. *Journal of Clinical and Experimental Neuropsychology, 22,* 325–338.

Unverzagt, F. W., Hall, K. S., Torke, A. M., Rediger, J. D., Mercado, N., Gureje, O., et al. (1996). Effects of age, education, and gender on CERAD neuropsychological test performance in an African American sample. *The Clinical Neuropsychologist, 10,* 180–190.

Vanderploeg, R. D., Schinka, J. A., Jones, T., Small, B. J., Graves, A. B., & Mortimer, J. A. (2000). Elderly norms for the hopkins verbal learning test-revised. *The Clinical Neuropsychologist, 14*(3), 318–324.

Wagner, M. T., Wymer, J. H., Carlozzi, N. E., Bachman, D., Walker, A., & Mintzer, J. (2007). Preliminary examination of progression of Alzheimer's disease in a rural Southern African American cohort. *Archives of Clinical Neuropsychology, 22*(3), 405–414.

Warrington, E. K. (1984). Recognition memory Test Ô (RMTÔ). Test Booklet 1. NFER-Nelson Publishing Company.

Wechsler, D. (2001). Wechsler test of adult reading. San Antonio: The Psychological Corporation.

Wechsler, D. (2009). WMS-IV.: Wechsler memory scale-administration and scoring manual. Psychological Corporation.

Whitfield, K. E., Fillenbaum, G. G., Pieper, C., Albert, M. S., Berkman, L. F., Blazer, D. G., & Seeman, T. (2000). The effect of race and health-related factors on naming and memory the macarthur studies of successful aging. *Journal of Aging and Health, 12*(1), 69–89

Assessing Risk and Recidivism in African Americans

20

Natalie M. Bennett, Brian D. Leany
and Lorraine T. Benuto

Assessing Risk and Recidivism in African Americans

According to Carson and Sabol (2012), the total number of prisoners in state and federal correctional institutions in the USA was just under 1.6 million in the year 2011. Of these prisoners, about 38% were African American. Looking only at prisoners under state jurisdiction, about 40% of inmates convicted of violent crimes were African American. Table 20.1 includes a breakdown of the number of sentenced prisoners by type of violent crime and race.

The eventual release of many of these prisoners may be a concern for the safety of the general population. To address this concern, municipalities have developed various strategies designed to predict probabilities for risk of dangerousness and recidivism. To this end, a number of measures have been created with varying degrees of reliability and validity. When considering the reliability and validity of a measure, there is the question of how these psychometric properties apply to racial and ethnic minority populations, as many of the measures are developed using primarily Caucasian samples (Benuto and Leany 2014).

Clinical vs. Actuarial Prediction

An additional concern beyond the psychometrics of a test is the decision-making process itself. One strategy for decision-making is actuarial prediction. Actuarial prediction eliminates the human aspect of the decision and applies an empirically based formula in making decisions. An alternative strategy is deriving a clinical decision based upon the knowledge, experiences, and perspective of the clinician who, through personal training and expertise and exposure to clients and clinical settings, has developed an intuition about psychological decision-making. However, research has shown that clinical judgment alone tends to be less accurate than actuarial methods in general (Dawes et al. 1989). Some research has shown that actuarial methods tend to be more accurate than clinical judgment alone in risk assessment (e.g., Oleson et al. 2011). In the domain of risk assessment, the results of many empirical studies are much more complex than choosing one method over the other (Litwack 2001). Forensic decision-making can be especially difficult compared to decision-making in a nonforen-

N. Bennett (✉) · L. T. Benuto
Administrative Faculty, Department of Psychology
MS0296 University of Nevada, Reno, NV 89557, USA
e-mail: nmbennett06@gmail.com

B. D. Leany
Lake's Crossing Center, Maximum Security Facility
for Forensic Mental Health Services, 500 Galletti Way,
Sparks, NV 89504, USA
e-mail: leanyb@gmail.com

B. D. Leany · L. T. Benuto
Northcentral University, Core Faculty in the Graduate
School, 10000 E University Drive, Prescott Valley, AZ
86314, USA

L. T. Benuto
e-mail: dr.benuto@gmail.com

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans,*
DOI 10.1007/978-1-4939-1004-5_20, © Springer Science+Business Media New York 2015

Table 20.1 Number of sentenced prisoners under state jurisdiction in 2011, broken down by type of offense and race. (Source: Carson and Sabol 2012)

Type of violent offense	All inmates	Caucasian	African American	Hispanic
Murder	166,700	47,200	70,100	38,900
Manslaughter	21,500	8600	7800	3300
Rape	70,200	32,500	22,200	8600
Other sexual assault	90,600	44,100	17,300	26,200
Robbery	185,800	40,400	96,600	38,000
Assault	146,800	44,300	57,200	38,500
Other violent	43,400	14,900	15,400	10,700

Details may not add to total numbers due to rounding and missing data

sic setting due to the fact that criminal behavior is a low base-rate behavior (Smykla 1986). That is, the proportion of criminals in the normal population is small, and thus it is difficult to predict such behavior with accuracy. Some assessments currently in use integrate the clinical and actuarial methods (e.g., the Structured Assessment of Violence Risk in Youth (SAVRY); Borum et al. 2006). Consequently, although this chapter does discuss measures that include some degree of clinical decision-making, this chapter places greater emphasis on empirically based actuarial measures that have been scientifically validated to predict risk or recidivism.

Assessing Risk and Recidivism

The heading for this section is a bit misleading, as most of these measures are actually assessing risk within the construct of recidivism, rather than providing a prospective appraisal of risk. That is, the application of these assessments traditionally is requested subsequent to the commission of a crime, and is incorporated as a component of parole, probation, and sentencing proceedings. The measures identified here are not exhaustive, but rather represent the more common measures used for the assessment of risk for reoffending.

Hare Psychopathy Checklist-Revised (PCL-R)

While not specifically derived as a risk assessment measure (in fact the author provides strong caveats against doing so in the user's manual), the PCL-R (Hare 2003) is a checklist used to assess an individual's level of psychopathy. It is a 20-item scale that is designed to measure two factors: (1) Interpersonal and Affective Problems, and (2) Unstable and Antisocial Lifestyle. The PCL-R is a clinical judgment tool in which a clinician uses data collected from a semistructured interview as well as historical file information to assign ratings as to how well each item applies to the individual under evaluation. The results of this are then traditionally used to be a piece of information in the overall evaluation of risk of reoffending for the individual (e.g., if there is consistent information that suggests that the individual endorses symptoms and behaviors consistent with antisocial personality disorder and psychopathy, there is a greater probability of reoffending, based on the symptom criteria themselves).

Specific to African Americans, Cooke, Kosson, and Michie (2001) compared factor structure on scores on the PCL-R between Caucasian and African American participants. They found that the structure of psychopathy was the same in both groups. Furthermore, although small significant differences existed in the performance on specific items, no significant difference in test functioning was found. Thus, the authors concluded that the PCL-R is unbiased toward and can be used with African Americans.

Additionally, researchers (Sullivan et al. 2006) studied the reliability and construct validity of the PCL-R among male inmates of various ethnic backgrounds. They found minor differences in PCL-R scores among an African

American sample as compared to Latino and Caucasian samples. However, these differences did not affect the overall pattern of relationships in the measure. Therefore, the authors argued that the PCL-R is a valid and reliable measure of psychopathy in adult African American inmates.

The PCL-R was also adapted for use with adolescents (Hare Psychopathy Checklist: Youth Version; PCL: YV; Forth et al. 2003). However, the need for caution with African American youth may exist with the PCL: YV. Researchers (Edens et al. 2007) conducted a meta-analysis of the relationship between psychopathy measures and criminal recidivism. They found that the relationship between violent recidivism and psychopathy was moderated by ethnicity. That is, the data suggested that studies which had a higher proportion of non white juveniles in their sample had results that indicated a weaker relationship between violent recidivism and psychopathy.

Summary and Recommendations The PCL-R has been shown to be generally unbiased and can be used with African American adults. However, there is some evidence that results from the PCL-YV may be moderated by ethnicity, and not as predictive of violent recidivism for African American youth. Thus, caution should be used when interpreting the results of the PCL-YV when used with African Americans.

Classification of Violence Risk (COVR)

The COVR (Monahan et al. 2006) is a software program that was developed as an actuarial tool to aid clinicians in estimating the risk that an individual hospitalized for a mental disorder will be violent to other people after discharge from the hospital. While not computer administrated, the software requires that the clinician enters information gathered from an interview with the individual, and compares this information to the test's normative data in order to generate a report that estimates an individual's risk for violence.

Initial search strategies revealed no specific studies on how African Americans perform on the COVR. Research in the last five years has shown that the COVR has predictive validity in general, although the samples used in such studies were mostly Caucasian. For example, McDermott, Dualan, and Scott (2011) found that the COVR was useful in predicting incidents of physical aggression among 146 patients in a psychiatric facility. About 44 % of this sample was non-Caucasian, but additional racial breakdowns were not provided. Additionally, other research has shown that the COVR was a good predictor of verbal and physical aggression among a sample of 52 participants in a psychiatric inpatient facility (Snowden et al. 2009); however, only 8 % of this sample was non-Caucasian. Again, additional racial breakdowns were not provided.

Summary and Recommendations No studies found have empirically tested the COVR for racial or ethnic bias, but it does show predictive validity for populations in general. While the predictive ability of the measure is encouraging, the lack of specific racial and ethnic data warrants a recommendation to use this measure with a great deal of caution.

Historical-Clinical-Risk Management-20 (HCR-20)

The HCR-20 (Webster et al. 1997) is a 20-item assessment device that contains ten historical factor items, five clinical items, and five risk management items. Initial search strategies revealed no specific studies on how African Americans perform on the HCR-20. Research in the last 5 years has shown that the HCR-20 has predictive validity in general, although the samples used in such studies were not very diverse. For example, Gray, Taylor, and Snowden (2008) tested the predictive validity of the HCR-20 with 887 male patients in a psychiatric facility. Although the majority of participants in this study were Caucasian, 21.6 % were of Black Caribbean or Black African origin. The authors found that the measure effectively predicted violent and nonviolent offending after release from the facility. Additionally, this information can still be useful in generalizing to

African Americans, as the US population is made up of 13.1 % African Americans (United States Census Bureau 2012). Thus, a higher proportion of African Americans constitutes the study sample than the general population of the United States.

Summary and Recommendations No studies found have specifically tested the HCR-20 for racial or ethnic bias, but it does show predictive validity for populations in general. Further, there are several studies that have overrepresentative samples of African Americans and Blacks, without indication of racial or ethnic bias. Further, even if the clinician were to forego this as a risk appraisal measure, the additional information related to clinical and management items, makes it a critical tool for providing recommended interventions to reduce risk, and warrants consideration in that capacity.

Level of Service Inventory-Revised (LSI-R)

An example of a semistructured interview designed to predict risk, the LSI-R (Andrews and Bonta 2001) is a quantitative assessment device that measures an individual's risk of reconviction and, similar to the HCR-20, identifies problem areas and needs, assessing both static and dynamic risk factors. The resulting responses yield scores for ten scales that include: Criminal History, Education/Employment, Financial, Family/Marital, Accommodation, Leisure/Recreation, Companions, Alcohol/Drug Problems, Emotional/Personal, and Attitudes/Orientation. The LSI-R has been shown to effectively predict recidivism in various populations (Vose et al. 2008).

In one study, the researchers (Schlager and Simourd 2007) examined the use of the LSI-R with African American and Hispanic populations, including 446 male parolees; 75 % of which were African American. The authors' concluded through their analyses that the LSI-R scores among their sample demonstrated similar psychometric properties when compared to other studies with Caucasian participants, although they noted that the predictive validity was lower yet still acceptable. Furthermore, the authors indicated that although slight differences in subcomponent variables emerged, these differences are likely to be clinically insignificant, and thus concluded that the LSI-R can be used effectively with African American offenders.

Interestingly, Whitacre (2006) tested the LSI-R at a federal community corrections center with 532 male residents. Using the LSI-R as a classification tool to discriminate between high risk and low risk for disciplinary incidents while incarcerated, as well as program outcome (i.e., successful completion or unsuccessful termination of the program), the author found that African Americans were more likely to be erroneously overclassified as high risk as well as under classified as low risk than either Caucasians or Hispanics. Whitacre thus cautioned that before implementing the use of the LSI-R as a classification tool in a correctional facility, one should conduct local validation studies to test for racial bias.

Summary and Recommendations The LSI-R has been shown to have acceptable predictive validity when used with African Americans; however, this validity was found to be lower in such populations when compared with Caucasian populations. Additionally, one study showed that the LSI-R resulted in more risk classification errors at both ends of the spectrum (e.g., overclassifying as well as underclassifying risk) for African Americans than for other populations. While the authors recommended local validation as a means to improve the predictive validity of the measure, this does not seem to be a practical solution for most forensic evaluators. Thus, results of the LSI-R should be interpreted with extreme caution when used with African American populations.

Violence Risk Appraisal Guide (VRAG)

Another example of a computer-aided assessment tool, the VRAG (Quinsey et al. 1998) was developed using a sample of 618 mentally disordered male offenders charged with a serious criminal offense. It is an actuarial tool that uses 12 weighted variables to predict the probability that an individual will commit a violent crime after release. These 12 risk variables were found through multiple regression analyses to correlate highly with reoffense, and include: psychopathy (this requires a PCL-R; PCL-SV, or PCL-YV score), separation from parents at an early age, injury to victim, schizophrenia, no marital history, difficulties in elementary school, female victim, prior conditional release failure, history of property offense, early onset of offending, history of alcohol abuse, and personality disorder diagnosis (Harris et al. 1993). Once the requested information is entered into the assessment instrument, individuals are assigned to a risk category ranging from one (lowest risk) to nine (highest risk), based on comparative norms. A primary benefit of this measure and its related measure for risk of sexual reoffending (discussed below), is that it is freely available on the internet, and allows direct, online entry.

While initial search strategies revealed no specific studies on how African Americans perform on the VRAG there is research that has demonstrated general predictive validity of the VRAG for general and violent recidivism in a German psychiatric institution (Kröner et al. 2007). However, this study provided no racial breakdown of its participants.

Summary and Recommendations No studies were found that tested the VRAG for racial or ethnic bias, although it has been shown to have predictive validity in general. Thus, some caution should be used when interpreting the results of the VRAG when used with African Americans.

Sexual Offender Risk Appraisal Guide (SORAG)

The SORAG (Quinsey et al. 1998) is a modification of the VRAG risk assessment device, and was designed specifically to predict the risk of sexual reoffending for male sex offenders. It contains 14 items (10 of which are also contained in the VRAG), all of which are static variables. Similar to the VRAG, the SORAG is an actuarial tool which weighs these variables and then assigns a level of risk and probability of reoffense to the individual.

Here too, initial search strategies failed to yield specific studies on how African Americans perform on the SORAG. Research in the last 5 years has shown that the SORAG has predictive validity in general. For example, researchers (Langton et al. 2007) evaluated the SORAG and found that it showed moderate levels of predictive accuracy for serious recidivism in a sample of 468 adult male sex offenders in a federal penitentiary. Racial and ethnic breakdown of the participants was not provided.

Summary and Recommendations No studies were found that tested the SORAG specifically for racial or ethnic bias, although this measure has been shown to have predictive validity in general. Thus, some caution should be used when interpreting the results of the SORAG when used with African Americans.

Structured Assessment of Violence Risk in Youth (SAVRY)

Thus, so far we have only briefly discussed juveniles as an aside to the PCL. Designed specifically to predict risk of violence in juveniles aged 12–18, the SAVRY (Borum et al. 2006) was developed using the structured clinical judgment model. This means that it integrates actuarial decision-making with clinical judgment. The measure is comprised of items identified as protective factors as well as risk factors that are both static (e.g., age at first offense) and dynamic (e.g., substance abuse). These risk factors are classified into three domains including: historical, social/contextual, and clinical/individual. Risk factors are assigned scores of low, moderate, or high, and protective factors are marked as either absent or present. A summary risk rating (low, moderate, or high) is then assigned, based on the clinician's professional judgment as informed by the scoring of relevant factors in the measure.

The investigators (Vincent et al. 2011) studied how racial differences may affect the predictive validity of the SAVRY for both violent and nonviolent rearrests after 5 years. The authors used a sample of 480 incarcerated male juvenile offenders of which 38.8 % were African American. Results showed that the SAVRY effectively predicted violent rearrests among youths regardless of racial and ethnic differences. However, the SAVRY total scores were unable to significantly predict nonviolent rearrests for Hispanic and African American youths. The authors proposed that the lack of predictive validity for racial minority youth for nonviolent rearrests may have been due to the disproportionate minority arrest rate in many states. As such, the authors hypothesized that if they had used conviction data rather than arrest data, such racial differences may have not appeared in the results or may have been less dramatic.

Summary and Recommendations The SAVRY has been shown to effectively predict violent arrests of all youths regardless of race or ethnicity. Given the traditional application of risk assessment evaluations (usually reserved for violent or sexual offenses), this measure may be appropriate for use. However, it may be less able to predict nonviolent arrests for minority populations, and thus when used for such purposes, results should be interpreted with extreme caution.

General Summary and Recommendations

In this chapter, we have discussed many of the common measures used to assess risk for reoffending. As mentioned, it is interesting to note that many of the measures discussed were not specifically researched with regards to performance among different ethnic and racial groups, especially in light of the disproportionate representation of African Americans incarcerated as compared to the number of African Americans who make up the US population. However, in spite of the specific research related to racial and ethnic characteristics of these measures, the majority of them included African Americans in their normative sample (with an occasional study containing an over representation of African American participants; Gray et al. 2008), and yielded good predictive validity. Of note, only the LSI-R demonstrated problematic psychometrics when used with African Americans, with these problems occurring at both ends of the spectrum, resulting in missing risk as well as overestimating risk. Thus, clinicians providing forensic evaluations of risk and recidivism should feel confident in their ability to identify and employ appropriate measures for their African American client. Further, while some measures, such as the HCR-20, do not describe even sample characteristics that include African Americans, they may still be used as a tool in making recommendations to reduce risk (e.g., if substance use is endorsed as present, it can be recommended as a target of treatment) rather than an appraisal of risk for reoffending. Therefore, one could combine measures of risk evaluation, not only to look for issues of convergence/divergence (e.g., Do both the freely available VRAG and the COVR predict risk of reoffending?), but also to provide guidance on potential areas of intervention. Again, the appraisal of risk is not foolproof, and is likely one of the most challenging forensic constructs. However, it is believed that the information above should provide reasonable guidance to the clinician in selecting measures that are appropriately valid for the clinician evaluation an African American client.

References

Andrews, D., & Bonta, J. (2001). *LSI-R: The Level of Service Inventory-Revised user's manual*. North Tonawanda: MultiHealth Systems.

Benuto, L. T. & Leany, B. (2014). *Guide to Psychological Assessment with African Americans*. Springer.

Borum, R., Bartel, P., & Forth, A. (2006). *Structured Assessment of Violence Risk in Youth (SAVRY)*. Lutz: Psychological Assessment Resources.

Carson, E. A., & Sabol, W. J. (2012). Prisoners in 2011. *Bureau of Justice Statistics Bulletin*, U.S. Department of Justice.

Cooke, D. J., Kosson, D. S., & Michie, C. (2001). Psychopathy and ethnicity: Structural, item, and test generalizability of the psychopathy checklist-revised

(PCL-R) in Caucasian and African-American participants. *Psychological Assessment, 13*(4), 531–542.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668–1674.

Edens, J. F., Campbell, J. S., & Weir, J. M. (2007). Youth psychopathy and criminal recidivism: A meta-analysis of the psychopathy checklist measures. *Law and Human Behavior, 31,* 53–75.

Forth, A. E., Kosson, D., & Hare, R. (2003). *The Hare psychopathy checklist: Youth version*. New York: Multi-Health Systems.

Gray, N. S., Taylor, J., & Snowden, R. J. (2008). Predicting violent reconvictions using the HCR-20. *British Journal of Psychiatry, 192,* 384–387.

Hare, R. D., (2003). *Manual for the revised psychopathy checklist* (2nd ed.). Toronto: Multi-Health Systems.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20,* 315–335.

Kröner, C., Stadtland, C., Eidt, M., & Nedopil, N. (2007). The validity of the Violence Risk Appraisal Guide (VRAG) in predicting criminal recidivism. *Criminal Behavior and Mental Health, 17,* 89–100.

Langton, C. M., Barbaree, H. E., Seto, M. C., Peacock, E. J., Harkins, L., & Hansen, K. T. (2007). Actuarial assessment of risk for reoffense among adult sex offenders: Evaluating the predictive accuracy of the Static-2002 and five other instruments. *Criminal Justice and Behavior, 34,* 37–59.

Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy, and Law, 7*(2), 409–443.

McDermott, B. E., Dualan, I. V., & Scott, C. L. (2011). The predictive ability of the classification of violence risk (COVR) in a forensic psychiatric hospital. *Psychiatric Services, 62*(4), 430–433.

Monahan, J., Steadman, H. J., Appelbaum, P. S., Grisso, T., Mulvey, E. P., Roth, L. H., Robbins, P. C., Banks, S., & Silver, E. (2006). The classification of violence risk. *Behavioral Sciences and the Law, 24,* 721–730.

Oleson, J. C., VanBenschoten, S. W., Robinson, C. R., & Lowenkamp, C. T. (2011). Training to see risk: Measuring the accuracy of clinical and actuarial risk assessments among federal probation officers. *Federal Probation, 75*(2), 52–56.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Schlager, M. D., & Simourd, D. J. (2007). Validity of the Level of Service Inventory-Revised (LSI-R) among African American and Hispanic male offenders. *Criminal Justice and Behavior, 34,* 545–555.

Smykla, J. O. (1986). Critique concerning prediction in probation and parole: Some alternative suggestions. *International Journal of Offender Therapy and Comparative Criminology, 30,* 125–139.

Snowden, R. J., Gray, N. S., Taylor, J., & Fitzgerald, S. (2009). Assessing risk of future violence among forensic psychiatric inpatients with the classification of violence risk (COVR). *Psychiatric Services, 60*(11), 1522–1526.

Sullivan, E. A., Abramowitz, C. S., Lopez, M., & Kosson, D. S. (2006). Reliability and construct validity of the psychopathy checklist—revised for Latino, European American, and African American male inmates. *Psychological Assessment, 18*(4), 382–392.

United States Census Bureau. (2012). http://quickfacts.census.gov/qfd/states/00000.html. Accessed 25 Nov 2013.

Vincent, G. M., Chapman, J., & Cook, N. E. (2011). Risk-needs assessment in juvenile justice: Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior, 38*(1), 42–62.

Vose, B., Cullen, F., & Smith, P. (2008). The empirical status of the level of service inventory. *Federal Probation, 72*(3), 22–29.

Webster, C. D., Douglas, K. S., Eaves, D. & Hart, S. D. (1997). HCR-20: *Assessing risk for violence* (version 2). Vancouver: Simon Fraser University.

Whitacre, K. W. (2006). Testing the level of service inventory-revised (LSI-R) for racial/ethnic bias. *Criminal Justice Policy Review, 17*(3), 330–342.

Forensic Assessment with the African American Client

21

Lorraine T. Benuto, Brian D. Leany and Jena Garrick

Forensic Assessment with the African American Client

The field of forensic assessment is vast. In this chapter we focused our attention child custody evaluations, the assessment of competency in legal proceedings, and the assessment of capacity. We also reviewed the literature on the use of polygraphy with African American clients and offered a discussion on the assessment procedures in personal injury. Our review of the literature revealed that research specific to African Americans on the tests and measures discussed throughout this chapter is largely lacking. To this end, where appropriate, a theoretical perspective was taken and recommendations that were either empirically based or theoretically derived are provided. As the stakes in forensic assessment are high, it is important that results from psychological tests be compared to other sources of data (e.g., collateral contacts, record review, etc.) so that the results can either be confirmed or refuted. The clinician should also note weaknesses in the tests and interpret results with caution as needed.

Child Custody Evaluations

No ethnic group is exempt from divorce and parental disputes. In fact, specific to African Americans, the divorce rate has increased over the last 10 years and continues to rise (U.S. Census Bureau 2010). Thus, it remains important to understand factors that best help predict child adjustment during a divorce. Moreover, assessment domains that measure the family structure and promotion of the well-being of the child are essential (Emery 2011). Given the American Psychological Association's push for culturally sensitive psychological services (APA 2013), it is important to examine how standard assessment measures should be used with cultural minorities. In this section, we specifically examine evidence-based child custody evaluation procedures, and how these procedures can be used with African Americans.

As child custody evaluations require assessment of complex psychological factors (e.g., allegations of substance use, domestic violence, physical or sexual abuse, mental illness, etc.), mental health professionals are the preferred choice for conducting the said evaluations (Mason and Quirk 1997).

L. T. Benuto (✉)
Administrative Faculty, Department of Psychology
MS0296 University of Nevada, Reno, NV 89557, USA
e-mail: dr.benuto@gmail.com

B. D. Leany · L. T. Benuto
Northcentral University, Core Faculty in the Graduate
School, 10000 E University Drive, Prescott Valley, AZ
86314, USA

B. D. Leany
Lake's Crossing Center, Maximum Security Facility
for Forensic Mental Health Services, 500 Galletti Way,
Sparks, NV 89504, USA
e-mail: leanyb@gmail.com

J. Garrick
University of Nevada, Reno, 45 Llama Ct., Reno,
NV 89511, USA
e-mail: jena.garrick@gmail.com

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5_21,

Table 21.1 The egregious/promotive factors model

Factors						
Egregious factors	Parent-child relationship	Parenting skills	Environmental instability	Parent mental health	Excessive interparental conflict	
Promotive factors	Positive parenting	Parental school involvement	Promotion of interpersonal development	Promotion of mental health	Promotion of community involvement	Effective co-parenting

In addition to the child custody evaluations, psychologists may also be called upon to evaluate a parent (or caregiver's) capacity or ability to parent. While the legal ramifications of such an evaluation can be substantial (including a determination that the parent in question is not "fit" to parent), the assessment procedures for child custody evaluations and the evaluation of parental capacity do overlap to some degree. Therefore, the specific measures discussed within this section could also be utilized as part of a battery to evaluate parental capacity.

In the context of child custody evaluations, the best interest of the child is the dominant legal test, and although there can be wide variations in its interpretation by judges, it generally favors the custody arrangement that will best fulfill the needs of the children involved and foster their normative development (NRS 125.480; Herman and Bernet 1997). Despite the court's involvement, it is the clinicians who are called upon to conduct child custody evaluations to determine the best interest of the child and, to date only one set of evidence-based child custody guidelines has been developed. Specifically O'Donohue and Tolle (2012) outlined the promotive and risk factors that act as the most important determinants in terms of child outcome and development. They termed these guidelines as the Egregious/Promotive Factors Model (EPFM). As the crux of this book is to provide the reader with evidence-based guidelines, we recommend this model be used in child custody and parental capacity evaluations. Despite that, this model is not specific to the assessment of parental capacity, the measures discussed herein to assess each domain of the EPFM can also be utilized in the context of parental capacity assessments as can be the cultural recommendations described throughout.

The EPFM (O'Donohue and Tolle 2012) measures promotive factors which are factors that enhance mental wellness and treatment outcomes and egregious factors which are risk factors that could be detrimental to the child's development (see Table 21.1 for a summary of the factors). The EPFM advocates joint custody as the primary custody arrangement as it results in the best child outcomes, with the child having regular, positive interactions with both parents (Tolle 2010). The EPFM allows for the clinician's discretion in picking assessment measures that will help them to rule in or out the various factors for each parent, and O'Donohue and Tolle (2012) created a text that lists recommended measures. Within this section, several of these measures (and some additional ones) are reviewed with specific information regarding their use with African Americans.

The African American Client

To obtain information about African Americans and child custody evaluations, three relevant databases (PSYCHinfo, PSYCHarticles, and EBSCOhost) were searched using several combinations of the following key terms; forensic assessment*, child custody, African American*, competency, and capacity. Unfortunately, published literature specific to child custody evaluations with African Americans was absent. Our search did yield many results—mostly relating directly to the psychometric properties of specific assessment domains, e.g., the MMPI-II, the Parental Behavior Inventory, etc. As indicated above, we recommend that assessment measures that correspond with the factors in the EPFM be chosen, as the EPFM is the only evidence-based child custody evaluation model. A summary of these measures and the corresponding factors can be found in Tables 21.2 and 21.3. As research on

Table 21.2 Egregious factors

EPFM factor assessed	Measure used to assess EPFM factor	Description of measure
Parent–child relationship	Parent Child Relationship Inventory (PCRI)	A 78-item self-report questionnaire that measures parents perceived receipt of emotional and practical support, communication, and involvement with the child (Gerard 1994)
	Parenting Stress Index-Short Form (PSI-SF)	A 36-item assessment that evaluates perceived parental stress (Reitman et al. 2002)
	Parenting Stress Index (PSI)	A 120-item self-report questionnaire used to evaluate stressors in the parent–child relationship (Solis and Abidin 1991)
Parenting skills	Parent behavior checklist	A 100-item checklist that measures the strengths and needs of parents with children aged 1–5. It contains 3 main subsections, Expectation, Nurture, and Discipline (Fox 1994)
	Parenting sense of competence scale	A 16-item scale that measures mothers' beliefs, values, and perceived skills regarding being a mother (Gibaud-Wallston and Wandersmann 1978)
	Parent authority questionnaire	A 30-item per parent questionnaire used to determine parenting styles: Authoritarian, Permissive, and Authoritative (Buri 1991)
Environmental instability	The home observation for measurement of the environment	Measures support and interactions that the child receives at home. This measurement focuses on the child's experience (Caldwell and Bradley 1984, Cladwell et al. 2003)
	Child abuse potential inventory	A 160-item questionnaire that estimates parental risk in suspected cases of child abuse (Milner 1986)
Excessive intraparental conflict	The children's perception of interparental conflict scale	A 9-subscale measurement that assesses a child's perceptions of marital conflict (CPIC; Grych et al. 1992)
	Multidimensional assessment of interparental conflict scale	A scale that obtains parents and children's reports of interparental conflict (MAIC; Tschann et al. 1999)

Table 21.3 Promotive factors

Positive parenting	Parent Behavior Inventory (PBI)	An inventory that measures parenting behavior of parents of preschool and school age children (Lovejoy et al.1999)
	Alabama Parenting Questionnaire (APQ)	A 42-item questionnaire that measures five major dimensions of parenting: positive involvement with children, supervision and monitoring, use of positive discipline techniques, consistency in the use of such discipline, and use of corporal punishment (Frick 1991)
Parental school involvement	The Family Involvement Questionnaire (FIQ)	A multidimensional scale that uses parents or primary caregiver self reports of involvement in a child's preschool, kindergarten, or first-grade schooling (Fantuzzo et al. 2000)
Promotion of interpersonal development	The Multidimensional Scale of Perceived Social Support (MSPSS)	A 24-item scale measuring self reported assessments of perceived social support from family, friends, and one other close social contact (Zimet et al. 1988)
	Functional Status Questionnaire (FSQ)	A 34-item brief questionnaire designed to assess a patient's psychological, physical, social, and role function abilities (Jette et al.1986)
Promotion of mental health	Functional Status Questionnaire (FSQ)	A 34-item questionnaire designed to assess a patient's psychological, physical, social, and role function abilities (Jette et al.1986)
	The Child's Report of Parental Behavior Inventory (CPBI)	A 192-item measure of children's perceptions of family discipline by both the mother and the father (Schaefer 1965)
Promotion of community involvement	The Children's Assessment of Participation and Enjoyment (CAPE)	A 55-item self-report measure of children's participation in extracurricular leisure and recreational activities (King et al. 2004)
Effective coparenting	Parenting Alliance Measure (PAM)	A 20-item self-report measure of the strength of the perceived alliances of parents of children of age group 1–19 years old (Konold and Abidin 2001)

African Americans and the measures associated with the EPFM was almost completely absent, a review of recent literature on each measure was conducted to determine the inclusion rates of African Americans in studies on these measures. Findings from this literature review can be found in Tables 21.4 and 21.5. It is important to note that for the most part, African Americans were included in these studies at the very least as a minority of participants and research suggesting that the use of these measures with African Americans was largely lacking. That being said, for some of the measures the empirical literature revealed that some cultural considerations should be made. A discussion of these follows.

Recommendations for Child Custody Evaluations with African Americans

Most of the research summarized in Tables 21.4 and 21.5 did not focus exclusively on African Americans or on examining between group differences on these measures. That being said, for the most part within these studies, there was not any evidence suggesting that using these measures with African Americans is problematic with a few caveats. These are discussed below.

First, it is important to note that African American families may display higher parenting stress (Wikoff and Sherradon 2012). This higher level of stress may be associated with low socioeconomic status and educational level (McNeil et al. 2001) Therefore, when factoring parental stress into the evaluation, the clinician may wish to explore whether or not the parent's stress level can be reduced (e.g., can the parent be referred to social services to improve his or her financial situation?). Moreover, because ethnicity is related to parenting style and stress, racial identity is factored into some measures e.g., Parenting Stress Index (Pinderhughes et al. 2000). This is important as cross-cultural studies do show differing parenting styles across ethnic groups that affect youth outcomes (Roche et al. 2007). Moreover, African American families often use a kin family style where friends and close relatives are largely incorporated in raising the children (Mano 2009), which could serve to increase or decrease parental stress depending on the circumstances (Raikes and Thompson 2005). In fact, both males and African Americans reported lower levels of perceived social support even though they had more familial figures involved in the upbringing of a child. As indicated above, social support systems such as extended kin relationships can foster stress; and therefore, be perceived as unsupportive falsely (Bruwer et al. 2008; Canty-Mitchell and Zimet 2000). Thus, where extended kin relationships are present, the clinician should not assume that they are helpful or positive and should instead assess the extent to which such relationships provide support. It is also important to note that it may be difficult for the clinician to ascertain parental stress level if multiple parental roles are involved (Fernandez 2011; Mano 2009), and the clinician may need to assess all caregivers (even non-parents) to determine the best interest of the child.

Second, assessing parent mental health should be tailored to the individual case. In most parental assessments, an MMPI-2 is administered (see the chapter of the textbook on self-report personality measures) and other measures may be administered to rule in or not potentially existing mental health conditions. Most importantly, mental health conditions that could interfere with effective parenting should be ruled out. This includes but is not limited to: substance-related problems; personality disorders, severe anxiety or depression, and psychosis. Within this book there are several chapters dedicated to the use of assessment measures with African Americans by individual disorder. The reader is encouraged to first administer the MMPI-2 and then to determine what other disorders need to be assessed for and further review the relevant chapter for specific measures. Third, with regard to the parent–child relationship, it is important that the clinician be aware that it may be necessary to assess the parent–child relationship through extended family and kin relationships i.e. aunts, grandmothers, rather than just the mother or the father (Mano 2009). The clinician should specifically examine parental social support, involvement in the child's life, communication, consistent

Table 21.4 Egregious factors and relevancy to African Americans

EPFM factor assessed	Measure used to assess EPFM factor	Percentage of African Americans included in recent research*	Recommended for use with this population?
Parent–child relationship	Parent Child Relationship Inventory PCRI	Tobin et al. 2011: *32%* Schroeder and Kelley 2010: *5%*	*Yes
	Parenting Stress Index-Short Form (PSI-SF)	Reitman et al. 2002: 85% Haskett et al. 2006: 68%	Yes
	Parenting Stress Index (PSI)	Clark et al. 1989: 79% Bluestone and Tamis-LeMonda 1999: 100%	Yes
Parenting skills	Parent Behavior Checklist (PBC)	Mano et al. 2009: 100%	Yes
	Parenting Sense of Competence Scale	Bor and Sanders 2004: 32% Ohan et al. 2000: minority unknown Rogers and Matthews 2004: N/A Gilmore and Cuskelley 2008: N/A	Yes
	Parent Authority Questionnaire (PAQ)	Shearin 2001: 100% Buri 1991: N/A Raval et al. 2013: N/A	Yes*
Environmental instability	The Home Observation for Measurement of the Environment (HOME)	Bradley et al. 2001: N/A	Yes
	Child Abuse Potential Inventory (CAI)	Miller et al. 1991: 100% Chaffin and Valle 2002: 9%	Yes
Excessive intraparental conflict	The Children's Perception of Interparental Conflict Scale (CPIC)	Mcdonald and Grynch 2006: 60% minority Reese-Weber and Hessen-McInnis 2008: 14% unspecified Gonzalez et al. 2000: 8%	Yes

* means percentage was not indicated but population was included/mentioned in the study

Table 21.5 Promotive factors and relevancy to African Americans

EPFM factor assessed	Measure used to assess EPFM factor	Percentage of African Americans included in recent research	Recommended for use with this population?
Positive parenting	Multidimensional Assessment of Interparental Conflict Scale (MAIC)	Keurig and Simon 1996: 2 %	Yes
	Parent Behavior Inventory (PBI)	***	Yes
	Alabama Parenting Questionnaire (APQ)	Frick et al. 1999: 31 % Shelton et al. 1999: 25 %	Yes
Parental school involvement	The Family Involvement Questionnaire (FIQ)	Fantuzzo et al. 2000: 57 % Manz et al. 2004: 96 %	Yes
Promotion of interpersonal development	The Multidimensional Scale of Perceived Social Support (MSPSS)	Bruwer et al. 2008: 44 %	Yes
	Functional Status Questionnaire (FSQ)	***	Yes
Promotion of mental health	Functional Status Questionnaire (FSQ)	***	Yes
	The Child's Report of Parental Behavior Inventory (CPBI)	***	Yes
Promotion of community involvement	The Children's Assessment of Participation and Enjoyment (CAPE)	***	Yes
Effective coparenting	Parenting Alliance Measure (PAM)	Abidin and Brunner 1995: 19 % Konald and Abidin 2001:14.5 % Floyd et al. 1998: 3 % minority	Yes

discipline and limit setting, and acceptance of the child's growing autonomy (Kaufmann et al. 2000) as these factors predict child adjustment. Specific to African Americans, Mano (2009) noted that African American youth are more likely to develop emotional and behavioral problems and to receive treatment for psychiatric disorders than other ethnic groups. Per Mano, socioeconomic disadvantages such as low income, poverty, mental health conditions, and unemployment rates disproportionately affect African Americans. With a high divorce and/or separation rates among this population (U.S. Census Bureau), substantial socioeconomic disadvantages, and a kin style family structure (Mano 2009) African Americans have many unique features, and as such these factors should be considered within a child custody evaluation. Specifically, the clinician should still weigh the presence and/or absence of the egregious and promotive factors to develop some insight as to what the best interest of the child might be. This may or may not be having the child be in the custody of the father or mother, but also a kin such as an aunt/uncle, grandmother/father etc. Moreover, the clinician may wish to tailor his or her recommendations based on the challenges that one or both parents present with (e.g., if the child is exhibiting behavioral problems, the clinician should refer the family to a behavioral health specialist; if the family is experiencing financial hardship, a referral to social services may be appropriate etc.). Also, African Americans have been found to score lower on the Parent Child Relationship Inventory (PCRI) than Caucasians (Gerard 1994). Finally, measures such as the Family Involvement Questionnaire have been empirically validated and deemed culturally (Manz et al. 2000, 2004), although such measures do not allow for reports from multiple informants, which may not give a clear picture of an African American home.

Fourth, with regard to positive parenting some researchers have found that while African American parents may display a more authoritative parenting style, this may actually have some positive outcomes. Indeed on the Parental Authority Questionnaire, African Americans were observed to demonstrate higher rates of discipline and lower rates of physical affection. Even after controlling for education and SES, the rates of authoritativeness subscale scores seemed to be inflated compared to their Caucasian counterparts (Reitman et al. 2002). With regard to positive outcomes and authoritative parenting, Hill (1995) found that among adolescents, authoritative parenting was associated with positive characteristics such as cohesion, intellectual orientation, organization, and achievement in black families. Hill noted that while African American mothers had been touted as too controlling due to authoritative parenting approaches, that authoritative interactions are also filled with positive attributes such as expressiveness and good humor. Thus, when an African American parent is observed to display an authoritative style of parenting, it may be important to determine how this parenting style is functioning for the child before deciding where this factor should weigh in on the egregious vs. promotive spectrum.

Summary and Conclusions

A thorough review of the literature has shown that the assessment measures that are relevant to child custody evaluations have been normed on samples that did include African American although certainly this group has been underrepresented in the empirical literature (See Table 21.2 for a summary of the measures evaluated, the constructs reviewed, and their relevance to the African American client). As outlined above, there are a number of cultural considerations the clinician must make when determining the best interest of the child in the context of either a child custody evaluation or a parental capacity evaluation.

Assessing Competency in Legal Proceedings

The legal concept, competency, refers to the ability of an individual to understand legal proceedings and to stand trial, if found fit (Roesch and McLachlan 2010). In the case of incompetency, the adjustment or postponement of court proceedings is allowed so that the individual can participate effectively in the trial, and the integrity of the criminal justice system is kept (Roesch et al. 1999). The current standard for competency to stand trial in the United States was established in the court case Dusky v. United States (1960). In *Dusky*, it was held that:

> It is not enough for the district judge to find that "the defendant is oriented to time and place and has some recollection of events", but that the test must be whether he has sufficient present ability to consult with his lawyer with a reasonable degree of rational understanding—and whether he has a rational as well as factual understanding of the proceedings against him. (p. 402).

Mental health professionals indicated that Dusky is too ambiguous and individual states can set their own boundaries for competency and choose which measures they want to use—some of which are not heavily researched and empirically validated with all populations (Roesch et al. 1999; Favole 1983). Recently, Pirelli, Gottdiener, and Zapf (2011) examined 68 competency research studies to determine the most prevalent measures used. These measures are summarized in Table 21.6.

While historically, researchers (Reich and Wells 1985) noted that "relative to competent defendants, incompetent defendants tended to be black, unmarried, less educated, had more prior psychiatric hospitalizations, and were more likely to be psychotic" specific research on competency measures with African Americans is sparse. This being said while some of the historical notations made by Reich and Wells have upheld (e.g., defendants with a Psychotic Disorder, those who are unemployed, and those with a psychiatric hospitalization history are most likely to be found incompetent to stand trial) demographic factors such as ethnicity and gender have been determined to play a lesser role in competency (Pirelli et al. 2011). Moreover, the research on the commonly used measures to assess competency have included African Americans in their samples and most researchers (e.g., Roesch et al. 2006; Chinman 2003) have not identified data that suggests that competency measures are unfit for usage with African Americans although

Table 21.6 Competency measures

Assessment measure	Description
The Competency Screening Test (CST)	A 22-item test that measures an individual's ability to mentally prepare for trial, where each item is based on a factor within the legal definition of fitness for trial and the psychological conditions that may contribute to significant impairment of that ability (Lipsitt et al. 1971)
Georgia Court Competency Test (GCCT, GCCT-MSH)	A 21-item test that measures an individual's competency through visual location identification of key persons in a courtroom simulation. This test assesses the competent from individuals who need further evaluation (Wildman et al. 1978)
MacArthur Competence Assessment Tool Criminal Adjudication (MacCAT-CA)	A 22-item assessment tool that assesses the individuals understanding of the legal system. The basis of the test is a brief story and the answers to subsequent questions to decide if the questions needs to be re-administered (MacCAT-CA; Bonnie et al. 1996)
Interdisciplinary Fitness Interview (IFI, IFI-R)	A three-section measure covering: legal, psychopathological, and overall evaluation factors. It is designed to measure mental status issues as well as legal issues simultaneously (Golding et al. 1984, 1993)
Fitness Interview Test (FIT/FIT-R)	This test takes approximately 30 min to administer and focuses on the psycho legal abilities of a person; ability to communicate with counsel, understand proceedings etc. (FIT-R; Roesch et al. 1994)
Evaluation for Competency to Stand Trial-Revised (ECST-R)	An 18-item and 3 scales measure that assesses competency to stand trial, and 18 items, 5 scales for atypical presentation of symptoms. This measure can help detect feigned incompetency (Rogers et al. 1998)
Minnesota Multiphasic Personality Interview (MMPI/MMPI-2)	A clinical 10-scale instrument that assesses personality traits and psychopathology. The test is a true/false format that covers major categories of abnormal behavior such as depression, paranoia, etc. and assesses a client's problem solving skills, attitude and other factors (Butcher et al. 1989)
Wechsler Abbreviated Test of Intelligence (WASI)	A four-subtest measure using Vocabulary, Similarities, Block Design, and Matrix Reasoning, assesses intelligence reliably in adolescents and adults (Wechsler 1939, 2009, 2012)
Brief Psychiatric Rating Scale (BPRS)	An 18–24 item scale that assess psychiatric symptoms such as hallucinations, depression, elated mood, motor retardations, etc. (Gorham 1962)
Competency to Stand Trial Assessment Instrument (CAI)	A 13-item instrument that measures the basic understanding of legal concepts. The manual also includes examples of incapacity and suggested questions. (CAI; Laboratory of Community Psychiatry 1973)

considerations or minor alterations may be applied to some assessments.

In fact, there is a small body of research that suggests that some assessment measures (e.g., the CAI) have been demonstrated to detect incompetency more frequently with African Americans than other ethnic groups. As competency evaluations often incorporate IQ and education, SES and employment, marital status, and previous psychiatric hospitalization history, African Americans may be at a disadvantage (Nicholson and Kugler 1991). Albeit dated, Nicholson and Johnson (1991) examined relationship between scores on the Georgia Court Competency Test (GCCT) and the WAIS-R in a sample that was predominantly (65 %) African American. Results from this study indicated that the performance on the GCCT was positively correlated to the WAIS-R. The reliance on measures such as IQ, psycho legal ability, and diagnosis resulted in non-Whites performing poorer than their white counterparts on both assessments, the GCCT and the WAIS-R; however, when controlling for IQ the effect was mediated (Nicholson and Johnson 1991). Thus, while IQ can unduly influence the outcome for African Americans (more so than their Caucasian counterparts), there are means of accounting for this influence. More recently, a meta-analysis including 22 studies revealed that non-Whites are1.5 times more likely to be found incompetent (Pirelli et al. 2011). Specific to African Americans, in a longitudinal study, Yumoto (2008) assessed African American youth with a family background of alcoholism to measure validity of the Competency to Stand Trial Assessment Instrument (CAI). Convergent validity was high between dissociative symptoms and disorganized attachment, while discriminant validity was much lower between IQ and age with regard to insecure and secure attachments. These results suggest that at least with adolescents the instrument's detection of competency is much more dependent on factors external to competency, and can be unduly influenced by attachment. Thus, this measure is likely to be sufficient in the presence of clear psychotic symptoms or disorganized attachment. However, with more subtle symptoms and the more prevalent presentation of either secure or insecure attachment, it is preferable to use.

Summary and Recommendations for Assessing CST with African American Clients

As indicated above, research specifically examining how African Americans perform on CST evaluations is sorely lacking. What little research does exist suggests that factors that typically are included in CST evaluations that predict incompetency (e.g., psychotic disorders) may occur at a greater frequency among African Americans. This issue is most problematic for those cases that are marginal, and the solution to this issue is not readily apparent. An error of decision-making in either direction can have negative consequences. For example, a finding of competence, when an individual is incompetent, can negatively impact their ability to defend their case and receive a fair trial. Conversely, a finding of incompetence when an individual is competent can unnecessarily delay both the defendant's and the victim's due process. In these marginal cases, the best option is to utilize available clinical as well as competency related measures, and attempt to identify collateral information, in particular historical educational or clinical data (e.g. school records of special education, records of psychiatric hospitalization, prior criminal records) in order to formulate a comprehensive opinion. Finally, it is important to note that the general standard for competency is a preponderance of the evidence, and the determination of competency is one that is ultimately made by the presiding judge. Thus, while the determination of competency is an imperfect one, there exists a system of due process to try and accommodate this imprecise process.

Psychological Assessment of Capacity

Capacity is recognized as a socio-legal concept involving the ability of an individual to understand available choices, realize consequences, and ultimately follow through with their choices (Newberry and Patchet 2008). According to Benuto (2012), historically capacity was described as a global capacity whereas now capacity is conceptualized in terms of specific capacities—the focus on specific functional capacities with means of increasing them.

In 2008, the American Bar Association (ABA) and American psychological Association (APA) presented the *Assessment of Older Adults With Diminished Capacities: A Handbook for Psychologists* in which six main capacities (medical consent, sexual consent, financial, testamentary, driving, and independent living capacities) were identified. The ABA and the APA also provided suggestions for assessing functional elements, diagnoses, psychiatric or emotional factors, cognitive underpinnings, values and preferences, risk of harm, and level of supervision. While these guidelines are outlined for older adults, their usage can be applied to general capacity evaluations of diminished capacity as well. In the chart below, the guidelines as listed by the ABA and APA (2008) are discussed, and implications or considerations for African American clients are discussed.

Specific to African Americans, a number of cultural considerations should be made. As the majority of these cultural considerations are related to cognitive testing (either IQ or neuropsychological) or the presence of certain diagnoses, the reader is encouraged to review the cultural considerations made in the associated chapters of this book. Specifically, the chapters on dementia, assessing IQ, and neuropsychological assessment of this text provide useful recommendations on the cultural considerations that need to be made when assessing cognitive abilities in African Americans. In addition to the above, it is important to note that when assessing Activities of Daily Living (ADLs) and Instrumental Activities of Daily Living (IADLs) several cultural considerations also need to be made.

Specifically, there are a number of factors that are specific to African Americans that can impact ADL and IADL functioning including high prevalence of medical and psychological conditions (Thorpe et al. 2009; Nies et al. 1999) and sociodemographic factors (familial demands may also make it difficult for African Americans to engage in self-care: Carthron et al. 2010). For example, higher levels of mental health diagnoses can also contribute to an inability to perform basic duties such as bathing, driving a car, or managing finances on one's own (Fitzpatrick et al. 1997). While these factors can cause diminished ADL and IADL functioning, it is important to note that if the client's functioning is impaired the impairment should still be taken into account when evaluating the person's capacity. Conducting clinical interviews, contacting collateral informants, or conducting assessments can be helpful in determining the extent to which diminished ADL or IADL functioning impacts the client's capacity.

Summary and Recommendations

The ABA and APA (2008) offer an effective set of guidelines for assessing capacity. These guidelines should be used when assessing capacity with African American clients. The guidelines make a point to distinguish between clinical and legal competency, and determine how to recognize specific domains of functioning which is important for capacity evaluations. As the assessment of capacity involves the assessment of cognitive functioning, the clinician should refer to the chapters on dementia, IQ, and neuropsychological assessment of this text. They should also review any of the chapters specific to mental illness that are relevant to the mental health diagnoses that the client that might be impacting their capacity. The clinician should also assess for diminished ADL or IADL functioning, and determine the extent to which this impairment is impacting the person's ability to make decisions, etc. Lastly, as in good practice with any type of assessment, the clinician should not base any decision on a single measure. Instead, the clinician should use multiple means (e.g., clinical interview, collateral contact, record review, administration of psychological tests, etc.) to gather the information necessary to provide an opinion regarding the person's capacity.

Polygraphy

Polygraphy utilizes a person's physiological response as a means of deciphering if they are telling the truth. Baseline readings are obtained by asking blatantly true or false questions such as "Are you the president of the United States? Or, is your name…?" Common polygraph techniques are the relevanti–irrelevant technique, the comparison question technique, and the concealed information test (Matte 2012). There is some research supporting the use of polygraphs in identifying lies, but the evidence of its effectiveness is scarce and often demonstrates the need to empirically validate the polygraphs usage (O'Donohue et al. 2004; Folley et al. 2000). As the polygraph measures a physiological response, the ethnic differences in response to stress may be worth considering (Soto et al. 2012). Specifically, researchers have found that African Americans show greater vascular responsiveness, less myocardial responsiveness, and greater anxiety as a baseline during stressing tasks (Soto et al. 2012). This could theoretically impact the results of a polygraph, and therefore results yielded from stress-induced tasks may need to be interpreted with caution.

Personal Injury

Personal injury refers specifically to the physical harm of a person. Litigation refers to damages to property such as cars or houses. Personal injury

Table 21.7 Competency to stand trial (CST) assessment measures

CST assessment measure	Percentage of African Americans included in sample	Recommended for use with this population?
Competency to Stand Trial Assessment Instrument (CAI)	Cooper et al. 2003: 58 %	Yes*
Interdisciplinary Fitness Interview-Revised (IFI-R)	***	Yes
Fitness Interview Test (FIT-R)	***	Yes
Georgia Court Competency Test (GCCT)	Bagby et al. 1992: 63.6 % Ustad et al. 1996: 50 %	Yes*
MacArthur Competence Assessment Tool-Criminal Adjudication (MacCAT-CA)	Otto et al. 1998: 65 % (non-White)	Yes
Competency Screening Test (CST)	Nottingham and Mattson 1981: 28 % Ustad et al. 1996: 50 %	Yes* (with all populations there is a high false positive rate and inconsistent factor structure (Melton et al. 2007))

* means percentage was not indicated but population was included/mentioned in the study

evaluations can encompass workplace harassment, motor vehicle accidents, malpractice, physical assault, etc (Witt and Weitz 2007). A search using "Personal Injury", "Assessment", and "African Americans" did not yield any relevant results regarding guidelines or specific assessment protocols involving the African American client. Nonetheless, personal injury evaluations are relevant to African Americans considering high rates of work place discrimination (actual and perceived), sexual harassment, and physical injury (Rospenda et al. 2009).

Normative Standard vs. Self-Standard

When assessing for damages in personal injury, evaluations assessments can be made using the normative standard and the self-standard. The normative standard compares the client's scores to those of others who are similar to them in some way (i.e., demographic characteristics)—essentially the goal is to compare functioning to a normative group standard. The self-standard compares the clients functioning pre and post injury to determine damages (Lanham and Misukanis 1999). Regardless of the approach for conducting the assessment, most typical assessment procedures include a clinical interview, collateral contacts, and the administration of psychological tests. As each of these is covered extensively in other parts of this book, we will not review these extensively herein, and recommend that the clinician utilize the cultural considerations made in other chapters of this book that align with the specific measures.

Conclusions

As evidenced in this chapter, research specific to African Americans on the tests and measures discussed throughout is largely lacking. In many instances, it was difficult, if not impossible, to provide purely empirical recommendations based on large quantities of data. Where this was the case, a theoretical perspective was taken and appropriate recommendations were provided. Perhaps, the most salient recommendation that exists is that when completing a forensic evaluation with a cultural minority client it is of utmost importance that test results be compared to other sources of data (e.g., collateral contacts, record review etc.) so that the results can either be confirmed or refuted. Where there is a discrepancy between test results and other sources of data, additional data should be sought (e.g., additional collateral contacts, the administration of additional tests) so that a clearer picture of the client's functioning can be obtained. The clinician should note weaknesses in tests and interpret results with caution as deemed appropriate, and the considerations discussed throughout this chapter should be made (Tables 21.7 and 21.8).

Table 21.8 Capacity evaluation factors and guidelines

Factor	Examples	Recommended assessment measures/approaches
Activities of Daily Living (ADLs) and Instrumental Activities of Daily Living (IADLs)	(ADL) grooming, toileting, eating, transferring, dressing (IADLs) ability to manage finances, taking medicine, and functioning in the home and community-grocery shopping, telephoning	All assessments listed in the competency section of the chapter are relevant to measure ADLs and IADLs
Diagnoses	Current or past emotional or psychiatric diagnoses	Record review (medical, psychological etc.), collateral contact, clinical interviews, and symptom-specific assessment measures
Cognitive underpinnings	Specific examples listed below in bold	Depends on the capacity to be assessed. Specifics listed below
Attentional abilities	Ability to attend to objects or tasks	Digit span, coding
Language	Speech production, language comprehension, and written language skills	Obtain language samples through scene description or having client following basic commands. Showing a line picture and asking for the correct name can assess object naming. Asking for written samples or having client answer questions after reading a passage shows comprehensions skills. Impairment could be indicative of dementia or aphasia
Memory	Recall, retention, learn new information	The Wechsler Memory Scales-IV, California Learning Test, MMSE (See chapter on neuropsychological assessment and dementia)
Visual perceptual abilities	Driving, financial calculations	Test if client can copy figures, decipher or match patterns, and/or construct objects to samples
Processing speed	Can result in poor decision making skills, delayed speech, motor skills, or attention	Assess via several subtests in the WAIS (See chapter on IQ), Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) or Trails A from the trail-making test (see the chapter on dementia)
Executive functioning	Ability to plan, think flexibly, respond to feedback, and inhibit impulsive responses	MMSE, RBANS, and COGNISTAT
Judgment and reasoning	Ability to make sound decisions, demonstrate capacity	The Similarities subtest of the WAIS, and Kaplan Practical Problem Solving Task, or reasoning from the COGNISTAT

References

Abidin, R. R. (1995). *Parenting stress index: Professional manual*. Lutz: Psychological Assessment Resources.

Abidin, R. R., & Brunner, J. F. (1995). Development of a parenting alliance inventory. *Journal of Clinical Child Psychology, 24*(1), 31.

Bagby, R., Nicholson, R. A., Rogers, R., & Nussbaum, D. (1992). Domains of competency to stand trial: A factor analytic study. *Law and Human Behavior, 16*(5), 491–507. doi:10.1007/BF01044620.

Bluestone, C., & Tamis-LeMonda, C. S. (1999). Correlates of parenting style in predominantly working- and middleclass African American mothers. *Journal of Marriage and the Family, 61,* 881–893.

Bor, W., & Sanders, M. R. (2004). Correlates of self-reported coercive parenting of preschool-aged children at high risk for the development of conduct problems. *Australian and New Zealand Journal of Psychiatry, 38,* 738–745. doi:10.1111/j.1440-1614.2004.01452.x.

Bradley, R. H., & Caldwell, B. M. (1984). The home inventory and family demographics. *Developmental Psychology, 20*(2), 315.

Bradley, R. H., Corwyn, R. F., & McAdoo, H. (2001). The home environments of children in the United States. *Child Development, 72*(6), 1844–1867.

Bruwer, B., Emsley, R., Kidd, M., Lochner, C., & Seedat, S. (2008). Psychometric properties of the multidimensional scale of perceived social support in youth. *Comprehensive Psychiatry, 49*(2), 195–201.

Buri, J. R. (1991). Parental authority questionnaire. *Journal of Personality and Social Assessment, 57,* 110–119.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administration, scoring and interpretation* (revised). Minneapolis: University of Minnesota Press.

Caldwell, R. M., Mandracchia, S. A., Ross, S. A., & Silver, N. C. (2003). Competency to stand trial and criminal responsibility: An examination of racial and gender differences among African American and Caucasian pretrial defendants. *American Journal of Forensic Psychology, 21,* 5–19.

Canty-Mitchell, J., & Zimet, G. D. (2000). Psychometric properties of the multidimensional scale of perceived social support in urban adolescents. *American Journal of Community Psychology, 28*(3), 391–400.

Carthron, D. L., Johnson, T. M., Hubbart, T. D., Strickland, C., & Nance, K. (2010). Give me some sugar! The diabetes self-management activities of African-American primary caregiving grandmothers. *Journal of Nursing Scholarship, 42,* 330–337.

Chaffin, M., & Valle, L. A. (2002). Dynamic prediction characteristics of the child. *Child Abuse & Neglect, 27,* 463–481.

Chinman, M., Young, A. S., Rowe, M., Forquer, S., Knight, E., & Miller, A. (2003). An instrument to assess competencies of providers treating severe mental illness. *Mental Health Services Research, 5*(2), 97–108.

Cooper, V. G., & Zapf, P. A. (2003). Predictor variables in competency to stand trial decisions. *Law and Human Behavior, 27*(4), 423–436.

Dusky v. United States. (1960). 362 U.S. 402.

Emery, R. E. (2011). *Renegotiating family relationships: Divorce, child custody, and mediation*. NY: Guilford Press.

Fantuzzo, J., Tighe, E., & Childs, S. (2000). Family involvement questionnaire: A multivariate assessment of family participation in early childhood education. *Journal of Educational Psychology, 92*(2), 367–375.

Favole, R. J. (1983). Mental disability in the American criminal process: A four issue survey. In J. Monahan & H. J. Steadman (Eds.), *Mentally disordered offenders: Perspectives from law and social science* (pp. 247–295). New York: Plenum.

Fernandez, M. A., Butler, A. M., & Eyberg, S. M. (2011). Treatment outcome for low socioeconomic status African American families in parent-child interaction therapy: A pilot study. *Child & Family Behavior Therapy, 33*(1), 32–48.

FIT-R. Fitness interview test-revised. A structured interview for assessing competency to stand trial. Sarasota: Professional Resource Press/Professional Resource Exchange.

Floyd, F. J., Gilliom, L. A., & Costigan, C. L. (1998). Marriage and the parenting alliance: Longitudinal prediction of change in parenting perceptions and behaviors. *Child Development,* 69(5), 1461–1479.

Fox, R. (1994). *Parent behavior checklist*.

Frick, P. J. (1991). The *Alabama parenting questionnaire.* Unpublished rating scale, University of Alabama.

Frick, P. J., Christian, R. E., & Wootton, J. M. (1999). Age trends in the association between parenting practices and conduct problems. *Behavior modification, 23*(1), 106–128.

Gerard, A. B. (1994). *Parent-child relationship inventory*. Los Angeles, CA: Western Psychological Services.

Gibaud-Wallston, J., & Wandersman, L. P. (1978). *Development and utility of the parenting sense of competence scale.* In: Paper presented at the meeting of the American Psychological Association, Toronto.

Gilmore, L., & Cuskelly, M. (2009). Factor structure of the parenting sense of competence scale using a normative sample. *Child: Care, Health and Development, 35,* 48–55.

Golding, S. L. (1993). Interdisciplinary fitness interview-revised: A training manual. Unpublished manuscript.

Golding, S. L., Roesch, R., & Schreiber, J. (1984). Assessment and conceptualization of competency to stand trial: Preliminary data on the interdisciplinary fitness interview. *Law And Human Behavior, 8*(3–4), 321–334.

Gonzales, N. A., Pitts, S. C., Hill, N. E., & Roosa, M. W. (2000). A mediational model of the impact of interparental conflict on child adjustment in a multiethnic, low-income sample. *Journal of Family Psychology, 4*(3), 365–379.

Gorham, D. R. & Overall, J. E., (1962). The brief psychiatric rating scale. *Psychological reports, 10*(3), 799–812.

Grisso, T. (2003). *Evaluating competencies: Forensic assessments and instruments* (2nd ed.). New York: Kluwer Academic/Plenum Publishers.

Grych, J. H., Seid, M., & Fincham, F. D. (1992). Assessing marital conflict from the child's perspective: The children's perception of interparental conflict scale. *Child development, 63*(3), 558–572.

Haskett, M. E., Ahern, L. S., Ward, C. S., & Allaire, J. C. (2006). Factor structure and validity of the parenting stress index-short form. *Journal of Clinical Child and Adolescent Psychology, 35*(2), 302–312.

Herman, S. P., & Bernet, W. (1997). Practice parameters for child custody evaluation. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(10 Suppl.), 57S–68S.

Hill, N. E. (1995). The relationship between family environment and parenting style: A preliminary study of African American families. *The Journal of Black Psychology, 21*(4), 408–423.

Hoge, S. K., Bonnie, R. J., Poythress, N., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1997). The MacArthur adjudicative competence study: Development and validation of a research instrument. *Law and Human Behavior, 21*(2), 141–179.

Jette, A. M., Davies, A. R., Cleary, P. D., Calkins, D. R., Rubenstein, L. V., Fink, A., & Delbanco, T. L. (1986). The functional status questionnaire. *Journal of General Internal Medicine, 1*(3), 143–149.

Kaufmann, D., Gesten, E., Santa Lucia, R. C., Salcedo, O., Rendina-Gobioff, G., & Gadd, R. (2000). The relationship between parenting style and children's adjustment: The parents' perspective. *Journal of Child and family studies, 9*(2), 231–245.

Kerig, P. K. (1996). Assessing the links between interparental conflict and child adjustment. The conflicts and problem-solving scales. *Journal of Family Psychology, 10*(4), 454–473.

King, G., Law, M., King, S., Hurley, P., Hanna, S., Kertoy, M., Rosenbaum, P., & Young, N. (2004). *Children's assessment of participation and enjoyment (CAPE) and preferences for activities of children (PAC)*. San Antonio: Harcourt Assessment, Inc.

Konold, T. R., & Abidin, R. R. (2001). Parenting alliance: A multifactor perspective. *Parenting Alliance Measure Assessment, 8*(1), 47–65.

Lanham, R. A., & Misukanis, T. (1999). Estimating premorbid intelligence: Determining change in cognition following brain injury. *Brain Injury Source, 3*(3), 42–43.

Lipsitt, P. D., Lelos, D., & McGarry, L. A. (1971). Competency for trial: A screening instrument. *American Journal of Psychiatry, 128*(1), 105–109.

Lovejoy, M. C., Weis, R., O'Hare, E., & Rubin, E. C. (1999). Development and initial validation of the parent behavior inventory. *Psychological Assessment, 11*(4), 534–545.

Manz, P. H., Fantuzzo, J. W., & Power, T. J. (2004). Multidimensional assessment of family involvement among urban elementary students. *Journal of School Psychology*, 42(6), 461–475.

Mano, K. E. J., Davies, W. H., Klein-Tasman, B. P., & Adesso, V. J. (2009). Measurement equivalence of the child behavior checklist among parents of African American adolescents. *Journal of Child and Family Studies, 18*(5), 606–620.

Matte, J. A. (2012). Critique of Meta-Analytic Survey of Criterion Accuracy of Validated Polygraph Techniques. Polygraph, 19.

McDonald, R., & Grynch, J. H. (2006). Young children's appraisals of interparental conflict: Measurement and links with adjustment problems. *Journal of Family Psychology, 20*(1), 88–99.

McNeil, C., Capage, L., & Bennett, G. (2001). Cultural issues in the treatment of young African American children diagnosed with disruptive behavior disorders. *Pediatric Psychology, 27*(4), 339–350.

Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (Eds.). (2007). *Psychological evaluations for the courts*: *A handbook for mental health professionals and lawyers (3rd ed.)*. New York, NY: The Guilford Press.

Miller, T. R., Handal, P. J., Gilner, F. H., & Cross, J. F. (1991). The relationship of abuse and witnessing violence on the child abuse potential inventory with Black adolescents. *Journal of Family Violence, 6,* 351–363.

Milner, J. S. (1986). *The child abuse potential inventory: Manual* (2nd ed.). Webster: Psytec.

Nam, Y., Wikoff, N., & Sherraden, M. (2013). Racial and ethnic differences in parenting stress: Evidence from a statewide sample of New Mothers. *Journal of Child and Family Studies,* 1–11.

Newberry, A. M., & Pachet, A. K. (2008). An innovative framework for psychosocial assessment in complex mental capacity evaluations. *Psychology, Health & Medicine, 13*(4), 438–449.

Nicholson, R. A., & Johnson, W. G. (1991). Prediction of competency to stand trial: Contribution of demographics, type of offense, clinical characteristics, and psycholegal ability. *International Journal of Law and Psychiatry, 14*(3), 287–329.

Nicholson, R. A., & Kugler, K. E. (1991). Competent and incompetent criminal defendants: A quantitative review of comparative research. *Psychological Bulletin, 109*(3), 355–370.

Nies, M. A., Vollman, M., & Cook, T. (1999). African American women's experiences with physical activity in their daily lives. *Public Health Nursing, 16,* 23–36. doi:10.1046/j.1525-1446.1999.00023.x.

Nottingham, E., & Mattson, R. E. (1981). A validation study of the competency screening test. *Law And Human Behavior, 5*(4), 329–335. doi:10.1007/BF01044948.

O'Donohue, W. T., & Levensky. E. R. (2004). Handbook of forensic psychology: Resource for mental health and legal professionals. Amsterdam: Elsevier Academic Press.

O'Donohue, W. & Woodward Tolle, L. (2012). Improving the quality of child custody evaluations, a systematic model.

Ohan, J. L., Leung, D. W., & Johnston, C. (2000). The parenting sense of competence scale: Evidence of a stable factor structure and validity. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 32*(4), 251–261.

Otto, R. K., Poythress, N. G., Norman, G., Nicholson, R. A., Edens, J. F., Monahan, J., Bonnie, R. J., Hoge, S. K., & Eisenberg, M. (1998). Psychometric properties of the MacArthur competence assessment tool–criminal adjudication. *Psychological Assessment, 10*(4), 435–443. doi:10.1037/1040-3590.10.4.435.

Pinderhughes, E. E., Dodge, K. A., Bates, J. E., Pettit, G. S., & Zelli, A. (2000). Discipline responses: influences of parents' socioeconomic status, ethnicity, beliefs about parenting, stress, and cognitive-emotional processes. *Journal of family psychology, 14*(3), 380.

Pirelli, G., Gottdiener, W. H., & Zapf, P. A. (2011). A meta-analytic review of competency to stand trial research. *Psychology, Public Policy, and Law, 17*(1), 1–53.

Raikes, H. A., & Thompson, R. A. (2005). Efficacy and social support as predictors of parenting stress among families in poverty. *Infant Mental Health Journal, 26*(3), 177–190.

Raval, V. V., Ward, R., Raval, P. H., & Trivedi, S. S. (2013). Confirmatory and exploratory factor analyses of parental authority questionnaire in urban India. *Journal of Child and Family Studies, 22*(5), 707–718.

Reese-Weber, M., & Hesson-McInnis, M. (2008). The children's perception of interparental conflict scale: Comparing factor structures between developmental periods. *Educational and Psychological Measurement, 68*(6), 1008–1023.

Reich, J., & Wells, J. (1985). Psychiatric diagnosis and competency to stand trial. *Comprehensive Psychiatry, 26,* 421–432.

Reitman, D., Currier, R. O., & Stickle, T. R. (2002). A critical evaluation of the parenting stress index-short form (PSI-SF) in a head start population. *Journal of Clinical Child and Adolescent Psychology, 31*(3), 384–392.

Roesch, R., Ogloff, J. R., & Golding, S. L. (1994). Competency to stand trial: Legal and clinical issues. *Applied and Preventive Psychology, 2*(1), 43–51.

Roesch, R., Zapf, P. A., Golding, S. L., & Skeem, J. L. (1999). Defining and assessing competency to stand trial. *The handbook of forensic psychology, 2,* 327–349.

Roesch, R., Zapf, P. A., & Eaves, D. (2006). *Fitness interview test—revised: A structured interview for assessing competency to stand trial.* Sarasota, FL: Professional Resource Press.

Roesch, R., & McLachlan, K. (2010). Competency to stand trial. *Corsini Encyclopedia of Psychology*.

Rogers, H., & Matthews, J. (2004). The parenting sense of competence scale: Investigation of the factor structure, reliability, and validity for an Australian sample. *Australian Psychologist, 39*(1), 88–96.

Rogers, R., Tillbrook, C. E., & Sewell, K. W. (2004). Evaluation of competency to stand trial-revised (ECST-R). Odessa, FL: Psychological Assessment Resources.

Rospenda, K. M., Richman, J. A., & Shannon, C. A. (2009). Prevalence and mental health correlates of harassment and discrimination in the workplace: Results from a national study. *Journal of Interpersonal Violence, 24*(5), 819–843.

Saab, P. G., Llabre, M. M., Hurwitz, B. E., Frame, C. A., Reineke, L. J., Fins, A. I., McCalla, J., Cieply, L. K., & Schneiderman, N. (1992). Myocardial and peripheral vascular responses to behavioral challenges and their stability in Black and White Americans. *Psychophysiology, 29,* 384–397. doi:10.1111/j.1469-8986.1992.tb01712.x.

Schaefer, E. S. (1965). Children's reports of parental behavior: An inventory. *Child development,* 413–424.

Schreiber, J., Roesch, R., & Golding, S. (1987). An evaluation of procedures for assessing competency to stand trial. *Bulletin of The American Academy of Psychiatry & The Law, 15*(2), 187–203.

Schroeder, V. M., & Kelley, M. L. (2010). Family environment and parent–child relationships as related to executive functioning in children. *Early Child Development and Care, 180*(10), 1285–1298.

Shearin, S. A. (2001). A descriptive study of family environmental factors and academic achievement of African American middle school males, *62*(4), 1585–N/A.

Simons, L., Brody, G. R., Murry, V. M., Granberg, E., Chen, Y., Simons, R. L., & Gibbons, F. X. (2008). Differences between European Americans and African Americans in the association between child obesity and disrupted parenting. *Journal Of Comparative Family Studies, 39*(2), 207–228.

Solis, M. L., & Abidin, R. R. (1991). The Spanish version parenting stress index: A psychometric study. *Journal of Clinical Child and Adolescent Psychology, 20*(4), 372–378.

Soto, A. J., Roberts, N. A., Pole, N., Levenson, R. W., Burleson, M. H., King, A. R., & Breland-Noble, A. (2012). Elevated baseline anxiety among African Americans in laboratory research settings. *Journal of Psychophysiology, 26*(3), 105–115.

Stephen, P., Herman, M. D., & William Bernet, M. D. (1997). Summary of the practice parameters for child custody evaluation. *Journal of American Academy of Child Adolescent Psychiatry, 36*(12), 1784–1787.

Tobin, N. L., Seals, R. W., & Vincent, J. P. (2011). Response patterns on the parent–child relationship inventory in a simulated child custody evaluation. *Journal of Child Custody, 8*(4), 284–300.

Tolle, L. W., & O'Donohue, W. (2012). *Improving the quality of child custody evaluations: A systematic model.* New York, NY: Springer.

Tschann, J. M., Flores, E., Pasch, L. A., & Marin, B. V. (1999). Assessing interparental conflict: Reports of parents and adolescents in European American and Mexican American families. *Journal of Marriage and the Family,* 269–283.

Thorpe, R. J., Szanton, S. L., & Whitfield, K. (2009). Association between lung function and disability in African Americans. *Journal of Epidemiology and Community Health, 63*(7), 541–545.

U.S. Census Bureau. (2001). *America's families and living arrangements*, current population reports, *P20–537, 2001 and earlier reports*.

Ustad, K. L., Rogers, R., Sewell, K. W., & Guarnaccia, C. A. (1996). Restoration of competency to stand trial: Assessment with the Georgia Court competency test and the competency screening test. *Law And Human Behavior, 20*(2), 131–146. doi:10.1007/BF01499351.

Viljoen, J. L., Roesch, R., & Zapf, P. A. (2002). Interrater reliability of the fitness interview test across 4 professional groups. *The Canadian Journal Of Psychiatry/La Revue Canadienne De Psychiatrie, 47*(10), 945–952.

Warner, W. J. (2005). Polygraph testing: A utilitarian tool. *FBI Law Enforcement Bulletin, 74*(4), 10–13. http://search.proquest.com/docview/204133786?accountid=452.

Wechsler, D. (1939). *The measurement of adult intelligence* (p. 229). Baltimore: Williams & Witkins.

Wechsler, D. (2009). WMS-IV.: Wechsler memory scale-administrationand scoring manual. Psychological Corporation.

Wildman, R. W., Batchelor, E. S., Thompson, L., Nelson, F., Moore, J.,Patterson, M., & DeLaosa, M. (1978). *The Georgia court competency test: an attempt to develop a rapid quantitative measure for fitness for trial.* Milledgeville, GA: Forensic Services Division, Central State Hospital.

Witt, P. H., & Weitz, S. E. (2007). Personal injury evaluations in motor vehicle accident cases. *Journal of Psychiatry & Law, 35*(1), 3–24.

Yumoto, C. (2008). *Attachment representation in inner-city African American adolescents*. Dissertation Abstracts International, p. 68.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of personality assessment, 52*(1), 30–41.

Index

L. T. Benuto, B. D. Leany (eds.), *Guide to Psychological Assessment with African Americans*,
DOI 10.1007/978-1-4939-1004-5,