

Brandon K. Schultz, Clifton Mixon, Anne Dawson,
Craig Spiel, and Steven W. Evans

A vital component of school mental health (SMH) programming is *program evaluation*, which is the process of assessing how well SMH services meet local needs. SMH evaluators are interested in the value of a program while taking into account available resources and local goals, primarily to provide feedback to stakeholders (Wholey, Hatry, & Newcomer, 2010). But unfortunately, high-quality program evaluation is difficult to achieve (cf., de Anda, 2007; Rones & Hoagwood, 2000; Weist, 2005). In this chapter, we examine program evaluation from the standpoint of rural school practitioners (e.g., school counselors, school psychologists, consulting clinical psychologists). We contrast our topic with school-based research, which is commonly conducted by outside experts who investigate specific treatments or disabilities to advance a broader knowledge base. Program evaluations, on the other hand, are local audits that are typically not intended to speak beyond the evaluation setting.¹

¹It should be noted that some emerging research paradigms based on evaluator-researcher collaboration blur this distinction (e.g., practice-based evidence; Kratochwill et al., 2012).

B.K. Schultz (✉)
Department of Psychology, Eastern Carolina
University, Greenville, NC, USA
e-mail: SCHULTZB@ecu.edu

C. Mixon • A. Dawson • C. Spiel • S.W. Evans
College of Arts and Sciences, Ohio University,
Athens, OH, USA

SMH program evaluation introduces procedural challenges that are unlike those commonly encountered by school-based researchers. For instance, SMH referrals can reflect the entire spectrum of child mental health needs (Farmer, Burns, Phillips, Angold, & Costello, 2003), requiring evaluators to examine outcomes across a wide range of referral questions. Evaluators might also examine the impact of multiple, unrelated interventions to assess whether their combination reduces the need for costlier services (i.e., program accountability). For such reasons, evaluation is not synonymous with research, although one might inform the other.

If it is assumed that published SMH effectiveness studies indirectly reflect program evaluation, the findings would be a cause for concern. Even though most research studies in the schools are conducted by experienced researchers, outcomes are often mixed or even disappointing (e.g., Hoagwood & Erwin, 1997; Kimber, Sandell, & Bremberg, 2008; Peltonen, Qouta, Sarraj, & Punamäki, 2012; Watabe, Stewart, & Owens, 2013; Wei, Szumilas, & Kutcher, 2010).² Research conducted in rural schools is particularly discouraging due to methodological weaknesses and

²Our list of examples is far from exhaustive. Many effectiveness studies find statistically significant results, but the effect sizes are smaller than those found in efficacy trials. When taking into account that these studies are often well-resourced, it seems safe to conclude that real-world programs yield even smaller effects on average.

inconclusive findings (Arnold, Newman, Gaddy, & Dean, 2005). For such reasons, new efforts are underway to identify the factors that prevent efficacious treatments from working in real-world settings (i.e., implementation science) (Cook & Odom, 2013). Likely barriers include competing staff responsibilities, logistical issues, and a lack of educator support (Langley, Nadeem, Kataoka, Stein, & Jaycox, 2010), but clearly these same factors can complicate program evaluation as well. We believe the field is overdue for a revolution in program evaluation science to advance training and methodology in this misunderstood and often overlooked aspect of SMH.

In this chapter, we discuss professional standards and measurement concerns in the evaluation of rural SMH programs. We do not treat rurality solely in cultural terms because there is a lot of variability from region to region, but there are economic and organizational concerns that are important to consider in the context of program evaluation. Unlike their urban and suburban counterparts, rural practitioners cannot easily consolidate resources to achieve an economy of scale (cf., Chakraborty, Biswas, & Lewis, 2000), so a relatively large proportion of resources go to redundant services across schools. Whereas a large school district might pool resources, rural school teams often need to develop procedures independently and with fewer contributors. Thus, we focus our discussion on the components of program evaluation that rural evaluators working in small groups might find most challenging. In instances where program evaluation concerns apply more broadly, we point readers to several helpful resources and keep our comments brief. We then conclude this chapter by offering our thoughts for how to improve the evaluation of rural SMH programs based on over a decade of experience collaborating with diverse rural schools.

Standards for Program Evaluation

Program evaluation is not unique to school mental health. In fact, there is a rich history of program evaluation in many applied settings, resulting in widely accepted and readily applicable standards

of practice. Although a complete review of these standards is beyond the scope of this chapter (see Yarbrough, Shulha, Hopson, & Caruthers, 2010 for a comprehensive discussion), we examine utility, feasibility, and accuracy standards, with an emphasis on the challenges that can cause rural SMH evaluations to fall short.

Utility Standards

To meet the needs of stakeholders (e.g., teachers, parents, students, community leaders, taxpayers), it is vital to ensure that evaluation outcomes are relevant and relatable. Utility standards for program evaluation pertain to the usefulness of the evaluation data (Yarbrough et al., 2010). In the evaluation of mental health programs, stakeholders will clearly want to know whether the program reduced morbidity, and perhaps improved safety in the setting. However, in rural communities, mental illness generally has greater stigma than in urban and suburban settings (Hoyt, Conger, Valde, & Weihs, 1997; Jones, Cook, & Wang, 2011), so evaluation outcomes that emphasize mental health labels could discourage participation due to the relative lack of anonymity in small communities. At the same time, typical barriers to accessing child mental health care in rural communities, such as a lack of transportation, scarcity of community-based mental health care providers, and a lack of health insurance, might be circumvented by SMH services (Owens, Andrews, Collins, Griffeth, & Mahoney, 2011). Thus, the challenge is to communicate the potential benefits of a SMH program, and at the same time avoid issues that might discomfit families, educators, and other stakeholders.

In our view, rural SMH evaluation is strengthened when evaluators measure outcomes related to client functioning—particularly school-related functioning—while avoiding potentially stigmatizing mental health labels. Outcomes related to children's disorganization and attention problems, for example, provide information that is more helpful to rural families than a measure of "ADHD." Similarly, a program that improves grades by helping students overcome procrastination and

distraction may speak more effectively to rural parents than a program for adolescents who are “depressed.” By shifting away from stigmatized labels, program outcomes may become salient for rural families and educators (Owens et al., 2011).

Similarly, some stakeholders might expand the definition of SMH program success beyond academic or attendance outcomes to include student safety. School administrators and school board members in particular are often interested in seeing tangible reductions in student disciplinary actions. Although disciplinary actions predict long-term unwanted outcomes (e.g., Walker, Steiber, Ramsey, & O’Neill, 1993), these data can be problematic because schools rarely standardize the reporting systems. The lack of standardization may be a particular problem in rural schools (e.g., Michael et al., 2013), given the remoteness of some schools relative to others. In the interest of utility, evaluators might need to help teachers standardize the office referral process to provide valid and reliable answers to these questions over time. Without standardized definitions, stakeholders from different disciplines and backgrounds can have the experience of using different terms to describe similar concepts—a phenomenon that interdisciplinary teams often lament.

Feasibility Standards

Feasibility standards promote the efficiency of program evaluation, in part by assuring practicality and cost-effectiveness (Yarbrough et al., 2011). In rural settings, the feasibility of evaluation can be a concern due to limited resources. In our experience, there are limited options in rural settings for staffing and funding evaluation efforts. We have found that rural school administrators are hesitant to invest in new programs due to initial cost concerns, and then when existing programs become widely accepted, practitioners are hesitant to invest additional time and resources to evaluate their outcomes.

Beyond resource concerns, many SMH practitioners do not know how to demonstrate program effectiveness without expert help. School counselors, for example, receive little formal training

in program evaluation beyond service recording (e.g., the number of counseling sessions provided) (Astramovich, Coker, & Hoskins, 2005). At the same time, stakeholders have a right to know that SMH practitioners are meeting their fiduciary responsibilities as a condition of continued investment (Poirier & Osher, 2006). Thus, evaluation costs are as integral a component of SMH programs as staffing and material costs, even though evaluation costs can take resources away from service provision. For this reason, we believe that feasibility concerns pose some of the greatest challenges for rural SMH evaluation. For program evaluation efforts to advance it will be vital for evaluators to make use of affordable measurement tools (see Instrumentation section below); but perhaps more importantly, it will be vital to demonstrate the long-term cost savings when children receive effective mental health care in the schools. If evaluators can clearly demonstrate student benefit and cost-effectiveness, stakeholders in rural areas will likely find SMH initiatives desirable.

Accuracy Standards

Accuracy standards relate to the need to ensure that the conclusions based on the evaluation data are justified (Yarbrough et al., 2011). One of the most important considerations in evaluation accuracy is *treatment integrity*, defined as the degree to which programs are implemented as intended (Schulte, Easton, & Parker, 2009). Outcomes can be compromised and the quality of the outcome evaluation can be weakened when integrity is poor (Durlak & DuPre, 2008). Even high-quality outcomes evaluations reflect the effectiveness of the program as delivered rather than as it was intended. If, for example, a Daily Report Card intervention (DRC) is implemented only intermittently rather than on a daily basis, the impact is likely to be weakened (Owens, Murphy, Richerson, Girio, & Himawan, 2008), but it would be a mistake to conclude that the DRC is useless. Thus, we must evaluate the *process* of intervention, and treatment integrity measures speak to these concerns. Poor integrity

might explain why some efficacious programs fail when implemented in naturalistic school settings (Atkins, Fraizer, Adil, & Talbott, 2003).

Although published effectiveness studies in rural schools are rare, two findings in Appalachian public schools have implications for achieving evaluation accuracy. First, Owens et al. (2008) examined evidence-based practices in rural elementary schools using a naturalistic referral process rather than participant recruitment. The researchers included measures of treatment integrity in their study, including both dosage and adherence indicators. In terms of dosage, the number of sessions attended and the number of days the teachers collected intervention data were assessed. In terms of adherence, the clinicians rated how well the home-school component of the intervention included parents' adherence to home-based procedures. The results suggest that teachers and parents implemented the interventions with acceptable (albeit imperfect) integrity, thereby providing support for the accuracy of the results (Owens et al., 2008). In other words, the authors could safely conclude that their results spoke to the efficacy of the intervention as designed.

Second, Albright and Michael (2013) evaluated a SMH program (Assessment, Support, and Counseling (ASC); Albright et al., 2013; Michael, Wandler, & Quick, 2010) in rural high schools that also responded to real-world referrals rather than participant recruitment. To evaluate the accuracy of their results, the researchers recorded dosage variables similar to Owens and colleagues (e.g., number of sessions), but also included client ratings of the clinicians' effectiveness. Client ratings of the clinicians taken at treatment termination provided some evidence that the program was directly targeting client needs as they perceived them, which in this case were mostly crisis intervention needs. In effect, the client ratings spoke to the clinician's competency in addressing mental health needs, which is important to ensure client buy-in.

Taken together, these studies suggest that the accuracy of program evaluation can be improved through adherence and competency measures including service tracking, permanent products, and satisfaction ratings. Program evaluations would clearly benefit from integrating similar

measures, but other integrity strategies used in effectiveness trials may be less useful. For example, researchers often create manuals for their treatments to ensure adherence to a set protocol of interventions, but this may not always be a realistic option in referral-based SMH programs. Albright and Michael (2013), for instance, describe their treatments as including cognitive-behavioral therapy, crisis intervention, and school consultation, but due to the variety of referrals, manuals for every service proved infeasible. SMH program evaluators are likely to face this same challenge (see also Albright et al., 2013; Michael et al., 2013).

Assessment Strategies

Next we turn our attention to assessment strategies, which are the specific techniques used when collecting data during the evaluation process. Much is known about best practice assessment because the techniques are similar across treatment research and program evaluation. Still, given the challenges faced when evaluating SMH programs, it is vital to consider the standards outlined above. Even valid and reliable measures could have limited utility, for example, if the costs are too high or if the constructs measured do not convey meaningful results to stakeholders. For this reason, our assessment recommendations integrate the aforementioned evaluation standards—utility, feasibility, and accuracy standards—throughout.

To begin, it is important to think strategically about the purpose of a program evaluation. Aligning this purpose with the goals of stakeholders can be critical to assuring the relevance of the findings, but ensuring that the measures and methods selected adequately address the purpose will further increase the relevance of the evaluation. Errors in these initial decisions can be costly. For example, we are personally aware of a community mental health agency that was contracted to provide SMH services in a rural school district. As part of these efforts, the service provider was required to perform a program evaluation in the first year that was to be considered by the school board prior

to renewal for the following year. The service provider kept track of the number and length of sessions with students, as well as satisfaction data collected from teachers at the end of the year. At the board meeting, the service provider reported contact with large numbers of students at every school and data about the number of sessions per student. In addition, they provided quotes from teachers who said positive things about their services. Nevertheless, none of these data spoke to whether the services benefitted students, as was noted by one of the school board members. No data were collected from or about students, and the perspective of parents was ignored. Having nothing other than service records and teacher quotes upon which to base their decision, the school board cancelled the contract. We believe that these events highlight the importance of understanding the goals of the program and considering the choice of questions with the stakeholders in the program. As is clear in this example, the purpose of the evaluation guides the choice of outcome domains and sources of information, consistent with the aforementioned evaluation standards.

Outcome Domains

In a review of SMH effectiveness studies, Rones and Hoagwood (2000) identified three outcome domains that might prove useful for program evaluation depending on the purpose of the evaluation: symptoms, functional impairments, and service usage (e.g., special education).

Symptoms. Most psychiatric disorders and other conditions of interest to school mental health programs are identified and differentiated from each other by the presence of specific symptoms. Symptoms are the observable behavioral features of a disorder, and assessing change in symptoms has been one of the most prominent methods of evaluating outcomes from treatment of mental health problems among youth (Weisz, Doss, & Hawley, 2005). Symptoms also provide a means of identifying youths who present with subclinical manifestations of disorders (i.e., are at risk for developing a disorder) and may still benefit from school-based mental health pro-

grams (Angold, Costello, Farmer, Burns, & Erkanli, 1999). Although symptoms have been the most prominent means of evaluating outcomes among youth (Weisz et al., 2005), increasingly researchers are recognizing that narrowly focusing on symptom reduction is insufficient and that outcomes beyond symptoms are needed in order to contextualize results (refer to Utility Standards above).

Functional impairments. Another prominent outcome domain in the treatment research and program evaluation literatures involves the problems (or distress) in daily life caused by symptoms of a disorder, referred to as functional impairment. Impairment is a required feature for the diagnosis of all child and adult disorders in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association (APA), 2013) and represents the impact of the symptoms or disorder on the individual or others. It is often impairment, rather than symptoms of the disorder, that lead to the need for mental health services. In some cases, impairments may be a better predictor of adverse outcomes (e.g., school failure) than a formal diagnosis (Vander-Stoep, Weiss, McKnight, Beresford, & Cohen, 2002) and may therefore be a better measure of treatment outcomes than symptoms alone. Youth who receive mental health services often experience impairment in multiple domains, most notably in academic and social functioning. As we mentioned above, program evaluations can benefit by including measures of change in students' academic and social impairment over time. In terms of analysis, the *reliable change index* can tell the evaluator the degree to which each child's change over time exceeds typical variation, based on the test-retest reliability of the instrument (see Jacobson & Truax, 1991). Similar calculations can be applied to symptom measures, but stakeholders in schools are likely to find reductions in impairment—particularly, academic impairment—most compelling.

Service use. Another important indicator that may be of interest to stakeholders is the use of school or community services following program implementation. A recent report on children's mental health indicates that one in five children experience a mental health disorder each year,

resulting in an estimated \$247 billion cost in terms of treatment utilization, special education, juvenile justice services, and decreased productivity (Centers for Disease Control and Prevention, 2013). It is not surprising that many program evaluations measure use of school and community services related to discipline, special education, mental health, and the juvenile justice system (Rones & Hoagwood, 2000). For example, in their evaluation of Project ACHIEVE, a comprehensive school reform process focused on improving outcomes for at-risk and underachieving students, Knoff and Batsche (1994) measured the number of students referred for special education assessment and the number placed in special education as outcomes. Similar types of service-use measures could clearly be useful in local program evaluations as well, and even seem to be implied as an outcome measure by many tiered prevention models (cf., Walker & Gresham, 2013).

Sources of Information

Several sources of information might be considered when conducting SMH program evaluation, but it may be difficult to determine whose data are most useful, feasible, and accurate. At the outset, it is important to determine if teacher, parent, child, clinician, observation, or school records will meet evaluation standards, so we discuss the strengths and weaknesses of each of these sources in turn.

Teacher reports. Teacher reports seem vital to high-quality SMH program evaluations, given the need for assessing school-related impairments; however, teacher reports are generally more accurate for identifying or assessing *externalizing* rather than *internalizing problems*, due to the overt/covert nature of these domains (McMahon & Frick, 2005; Pelham, Fabiano, & Massetti, 2005; Silverman & Ollendick, 2005). For example, teachers may easily observe when a student is refusing to comply with requests or is out of his or her chair (externalizing problems), but not notice when a student feels hopeless or worried (internalizing problems). Thus, we generally recommend prioritizing teacher reports when monitoring externalized concerns and parent or student

self-reports (described below) when monitoring internalized concerns. In secondary schools, the use of teacher reports is further complicated by the fact that students typically work with multiple teachers during the school day, so it is not always clear which teacher(s) ought to be included. There is some evidence for systematic teacher bias, including the tendency for women and early service teachers to be more severe in their ratings of externalizing disorders than men or more experienced teachers (Schultz & Evans, 2012), so evaluators must use caution when choosing, or weighing disagreements among, several teacher reporters.

Parent reports. Professional recommendations for child and adolescent assessment include the use of parent report for *both* internalizing and externalizing disorders (Hunsley & Mash, 2007). Parent report is important to assessing for childhood problems because parents are typically involved in children's day-to-day lives, making them informed reporters. Parent reports are also important because children may not be reliable when reporting the temporal sequence of their problems (Klein, Dougherty, & Olino, 2005). Thus, collecting parent reports can greatly inform outcome assessments of SMH programs, but there are some limitations. For example, parent psychopathology can bias these reports and we can reasonably expect such concerns in many SMH cases. As a case in point, it has been demonstrated that mothers with depression over-report symptoms of ADHD (Pelham et al., 2005) and depression in children (Klein et al., 2005). As a result, program evaluation cannot safely rely on these reports alone, and teacher ratings might be used to confirm or supplement parent data.

Child self-report. In general, child and adolescent self-reports are less reliable for externalizing problems as compared with internalizing problems. Children and adolescents with behavior disorders often underestimate aggressive behaviors, symptoms, and overall impairment when rating themselves (McMahon & Frick, 2005; Pelham et al., 2005). But self-reports can be useful when assessing outcomes related to covert forms of conduct problems, such as drug use, risky sex, and dangerous driving behaviors. For

internalizing problems, such as anxiety and depression, child and adolescents self-report can be more valid than either teacher or parent report. One caution regarding self-report of internalizing problems is that some groups of children and adolescents (e.g., younger children, African American, and Hispanic American youths) may be more likely to minimize problems, a bias sometimes referred to as *social desirability* (Silverman & Ollendick, 2005).

Clinician report. In some instances, evaluators might be tempted to have clinicians and interventionists report their impressions of outcomes, particularly when other information is missing or otherwise unavailable. The potential for bias in these reports is obvious, particularly when evaluations affect resources, but there are instances where clinicians can provide context for other evaluation outcomes. For instance, clinician report of the therapeutic alliance can predict outcomes to some degree (Elvins & Green, 2008; McLeod, 2011). As such, clinician alliance ratings can lend support for treatment competence, consistent with accuracy standards (assuming adequate validity and reliability). Clinician reports have also been used to measure how well parents have followed through with aspects of intervention. For example, Owens et al. (2008) asked clinicians to rate parent adherence to a home-school collaborative intervention, but the accuracy of these data was unclear.

Multiple informants. Given the potential for source biases, multiple-informant assessment is widely recommended, but inter-informant disagreement is common (Achenbach, McConaughy, & Howell, 1987). SMH practitioners often grapple with how to integrate the differing reports, understanding that disagreement between raters does not necessarily mean that the reports are invalid. Differences between raters could be due to differences in tolerance for child and adolescent behaviors or differences in how students behave across situations. Thus, variations in report across informants could offer valuable insight into target behavior, but potential rater biases still need to be considered.

Several suggestions for integrating reports from multiple informants can be found in the lit-

erature (Klein et al., 2005). One strategy is the “or” rule, which assumes that a behavior is present if it is reported by any informant. Alternatively, there is the “and” rule, which requires at least two informants confirm an observation. A third approach that more closely resembles clinical practice is the “best estimate” strategy, which relies on clinical judgment to integrate varying reports from informants. Although the “best estimate” strategy can introduce clinician bias, there is some evidence to suggest that the reliability of this estimate can be high (e.g., Klein, Ouimette, Kelly, Ferro, & Riso, 1994).

Direct observation. Systematic direct observation (SDO) typically involves observing students in their normal environments (e.g., classrooms) to assess changes in behavior. SDO can lead to accurate and contextualized measurement of behaviors that cause impairment, but observations often require significant staff resources (i.e., training, time) and often have limited reliability and generalizability without repeated observations (Hintze & Matthews, 2004). These concerns can render direct observations infeasible for large numbers of students. Moreover, low-frequency behaviors (e.g., physical aggression) are poor targets for SDO because the observer may not see the behavior of interest. Still, SDO of academic (e.g., off-task or disruptive behavior) and social (e.g., withdrawal) behaviors can provide important information about changes in student behavior over time.

Existing school data. Existing school data (e.g., grades, disciplinary referrals, attendance, and gradebook data) are attractive because of availability, relevance, and freedom from biases caused by social desirability, parent psychopathology, or evaluator judgment. For these reasons, existing data can serve as a convenient and valid outcome measure of SMH programs in some instances. But there are limitations because these data generally do not specify the duration, frequency, or intensity of specific problems (Riley-Tillman, Kalberer, & Chafouleas, 2005). There can also be questions regarding the reliability and validity of existing school data. For example, grading policies can vary across teachers or school districts and grades are influenced

by external social influences such as poverty and other major life stressors. Given these limitations, we recommend that evaluators only include school data as part of a multi-method, multi-informant strategy. When multiple sources of information support the same conclusion, evaluators can have confidence in their outcomes.

Instrumentation

Researchers and program evaluators alike have searched for “generic” measures to assess outcomes for myriad referral questions (Schulte, 2010). Similarly, it would be convenient to identify a General Outcome Measure (GOM; Deno, Mirkin, & Chiang, 1982) for each outcome domain with strong psychometrics and relevance (Walker & Gresham, 2013), but so far generic measures and GOMs have proven elusive. Instead, there are multiple candidate instruments for the various outcomes of interest, each with strengths and weaknesses. The options and tradeoffs can be overwhelming, especially given that instrumentation interjects as much variance into outcome estimates as do the actual treatments (Wilson & Lipsey, 2001). Thus, instrumentation is a critical element of program evaluation that requires careful consideration.

At the outset, evaluators must develop hypotheses about which outcomes the program interventions are most likely to impact, and then select relevant measures for each outcome of interest. It can be useful to also consider possible distal outcomes and measure those as part of this process as well. Distal outcomes can include nontarget symptoms (e.g., reduction in child depressive symptoms after an intervention targeting ADHD), additional functioning measures (e.g., peer sociometric ratings after a social skills intervention), environmental impact measures (e.g., reduction in maternal depressive symptoms after a behavioral parent training intervention), and client satisfaction measures (e.g., parent satisfaction with group parent training). For this reason, treatment research utilizes 12 participant measures on average (Weisz et al., 2005), which is likely to prove unrealistic for rural SMH program evaluations.

But by expanding the scope of assessment beyond one or two outcomes, evaluators can gain a comprehensive understanding of treatment outcomes relative to the target concerns, as well as the child’s overall level of functioning (e.g., classroom functioning, family functioning).

Table 21.1 provides a brief overview of instruments that can be useful for SMH program evaluation. Given the limited resources available in rural communities, we highlight free instruments that are readily available online. Each instrument listed includes information regarding the class of instrument (e.g., rating scale), source (e.g., parent, teacher), domain (e.g., symptoms, impairment, satisfaction, classroom or family functioning), construct assessed (e.g., depression, academic performance), and age range for which it has been validated.³ Of course, many more instruments could be useful depending on the needs of evaluators, so readers are encouraged to refer to the treatment effectiveness and program evaluation literatures. Readers needing a more thorough review of the various classes of instruments than is provided here are encouraged to read one of the excellent published reviews (e.g., Pelham et al., 2005; Riley-Tillman et al., 2005). Below we offer brief overviews of some of the instruments highlighted in Table 21.1.

Symptoms. In the symptom domain, instruments generally fall into two categories: broadband and narrowband. Broadband scales measure a wide variety of behavior concerns, generally including both externalizing and internalizing symptoms, whereas narrowband ratings focus on specific concerns, such as anxiety or depression. In program evaluation, we would predict that broadband ratings would prove most useful for the reasons stated at the beginning of this chapter. For example, the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a 25-item behavioral rating scale that can be used to screen for or progress monitor problems in a number of domains. Parents and teachers are asked to rate the severity of emotional symptoms, conduct problems, symptoms of hyperactivity/

³Internet links are available at <http://www.oucirs.org/resources/educator&mhprofessional>.

Table 21.1 Publicly available measures of symptoms and impairment

Instrument	Class	Source	Domain	Construct	Age range
Academic Performance Rating Scale (APRS; DuPaul et al., 1991)	Rating scale	Teacher	Impairment	Academic performance	6–12
Autism Treatment Evaluation Checklist (ATEC; Rimland & Edelson, 1999)	Rating scale	Parent/Teacher/Clinician	Symptoms and impairment	Autism spectrum	5–12
Brief Problem Checklist (BPC; Chorpita et al., 2010)	Interview	Child/Parent	Symptoms	Global	7–13
Center for Epidemiological Studies Depression Scale for Children (CEC-DC; Weissman, Orvaschel, & Padian, 1980)	Rating scale	Child	Symptoms	Depression	6–17
The Child PTSD Symptom Scale (CPSS; Foa, Johnson, Feeny, & Treadwell, 2001)	Rating scale	Child	Symptoms	PTSD	8–18
Children's Global Assessment Scale (CGAS; Shaffer et al., 1983)	Rating scale	Clinician	Impairment	Global	4–16
Children's Yale-Brown Obsessive Scale (CY-BOCS; Scahill et al., 1997)	Rating scale/Interview	Parent/Child/Clinician	Symptoms	OCD	8–18
Classroom Performance Survey (CPS; Brady, Evans, Berlin, Bumford, & Kern, 2012)	Rating scale	Teacher	Impairment	Classroom behavior	Grades 9–12
Client Satisfaction Questionnaire (CSQ-8; Larsen, Attkisson, Hargreaves, & Nguyen, 1979)	Rating scale	Parent	Satisfaction	Satisfaction with treatment	All
Columbia Impairment Scale (CIS; Bird et al., 1993)	Rating scale	Child/Parent	Impairment	Global	9–17
Disruptive Behaviors Disorders Scale (DBD; Pelham et al., 1992)	Rating scale	Parent/Child/Teacher	Symptoms	DBD	5–16
Hamilton Anxiety Rating Scale (HAM-A; Hamilton, 1959)	Rating scale	Clinician	Symptoms	Anxiety	6–18
Homework Performance Questionnaire (HPQ; Power, Dombrowski, Watkins, Mautone, & Eagle, 2007)	Rating scale	Parent/Teacher	Impairment	Homework performance	Grades 1–8
Hopelessness Scale for Children (HSC; Kazdin, Rodgers, & Colbus, 1986)	Rating scale	Child	Symptoms	Hopelessness/Suicidality	6–13
Impairment Rating Scale (IRS; Fabiano et al., 2006)	Rating scale/narrative	Parent/Teacher	Impairment	Overall	4–12
Pediatric Symptom Checklist (PSC; Little, Murphy, Jellinek, Bishop, & Arnett, 1994)	Rating scale	Parent/Child	Symptoms	Global	3–18
Screen for Childhood Anxiety Related Disorders (SCARED; Birmaher et al., 1997)	Rating scale	Child/Parent	Symptoms	Anxiety	8–18

(continued)

Table 21.1 (continued)

Instrument	Class	Source	Domain	Construct	Age range
Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997)	Rating scale	Parent/Teacher	Symptoms	Global	3–16
Traumatic Events Screening Inventory (TESI-C; Ribbe, 1996)	Interview	Clinician	Symptoms	Exposure to trauma	6–18
Young Mania Rating Scale (YMRS; Young, Biggs, Ziegler, & Meyer, 1978)	Interview/Rating scale	Clinician/Parent	Symptoms	Mania	5–17
Youth Top Problems (TP; Weisz et al., 2011)	Interview/Rating scale	Child/Parent	Symptoms	Global	7–13

Note: Hyperlinks to all of the instruments in this table can be found at <http://www.oucirs.org/resources/educator&mhprofessional>

inattention, peer relationship problems, and pro-social behavior. As such, the SDQ could be a useful outcome measure suitable for program evaluation purposes.

It is also conceivable that narrowband scales could be useful in evaluation if these measures are targeted to specific referral questions. If the evaluation is stratified by the referral category, for example, narrowband scales might be used for each. There are several examples of narrowband rating scales that could be useful for these purposes. For example, the Disruptive Behavior Disorders Rating Scale (DBD; Pelham, Gnagy, Greenslade, & Milich, 1992) assesses symptoms of a number of externalizing disorders common in childhood and adolescence, including attention deficit/hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), and conduct disorder (CD). Currently the items on this instrument are similar to the diagnostic items found in the DSM.

Functional impairment. Similar to the measures of symptoms, instruments that assess functional impairment can focus on specific domains of impairment (e.g., academic, social) or global outcomes. For example, the Academic Performance Rating Scale (APRS; DuPaul, Rapport, & Perriello, 1991) is a measure completed by teachers to rate a child's academic performance over the past week across a number of subject domains and academic abilities. Sometimes it may be more useful to impact of treatment across multiple domains of functioning or on the overall functioning of a child. The Impairment Rating Scale (IRS; Fabiano et al., 2006) has parents and teachers rate the severity of impairment and need for treatment in multiple areas (e.g., relationships with peers, siblings, parents, teacher; academic performance; self-esteem; and overall) resulting to the child's presenting problems. Overall impairment can also be assessed using rating scales filled out by either the clinician (e.g., The Children's Global Assessment Scale (CGAS); Shaffer et al., 1983) or parent and child report (e.g., Columbia Impairment Scale (CIS); Bird, Shaffer, Fisher, & Gould, 1993).

Systemic outcomes. Beyond measures of symptoms and impairment, program evaluators may be interested in the impact SMH programs have on families and schools. For example, the Parenting Stress Scale (PSS; Berry & Jones, 1995) has been used to obtain parent report on the amount of stress they experience in their role. Although we do not include such instruments in Table 21.1, readers can refer to Pritchett et al. (2011) for an extensive list of family functioning measures that could prove useful in SMH program evaluations.

Conclusion

We have explored the goals of program evaluation and how those goals might be achieved for rural SMH programs. As we have shown, program evaluation is not synonymous with school-based research, which often has expert support, limited foci, and implications for the broader field. Program evaluation, by comparison, is intended to assess whether SMH programs meet the needs and expectations of a local community. Standards for program evaluation guide the practice, but rural evaluators are likely to encounter economic and organizational challenges along the way. We have highlighted several of these challenges in this chapter, but we cannot anticipate all potential difficulties; readers will need to consider the possible roadblocks in their setting when planning an evaluation. High-quality program evaluation requires that utility, feasibility, and accuracy standards are maintained.

Early in this chapter, we claimed that school mental health was overdue for a revolution in program evaluation science. It seems this need has been overlooked partly due to confusion between treatment research and program evaluation. Field-based treatment research has the laudable goal of establishing the effectiveness of a given treatment or program, but we should also recognize that even the most well-established techniques could fail because of setting-specific incompatibilities. That failure may or may not represent a threat to the "evidence-based" status of the treatment, but

it will most certainly have implications for the future of that program in that particular setting. Thus, best practices require that program evaluations are conducted in all localities, regardless of whether the treatments are “evidence-based” or not. Of course, this statement in itself is not revolutionary, but the implications for training are. It is clear from the literature that SMH professionals are largely unprepared to conduct high-quality program evaluations without external support and added resources. If SMH professionals assume that quality evaluations are only conducted by expert researchers, the true impact of local SMH programs will go unexamined. In our experiences collaborating with rural schools, this is certainly the case—SMH practitioners rarely have their programs evaluated. When programs are evaluated, practitioners seem to assume that convenient data, such as service records and client grades, are sufficient to meet the needs of the stakeholders. But as we have pointed out, such data are inadequate for meeting the standards of quality program evaluation. The solution will require fundamental changes to how SMH professionals approach and evaluate their practice.

References

- Achenbach, T., McConaughy, S., & Howell, C. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232.
- Albright, A., & Michael, K.D. (2013, October). The effectiveness of a rural school mental health program: The assessment, support, and counseling center. *Paper presented at the 17th annual Conference on Advancing School Mental Health, Washington, DC.*
- Albright, A., Michael, K. D., Massey, C., Sale, R., Kirk, A., & Egan, T. (2013). An evaluation of an interdisciplinary rural school mental health programme in Appalachia. *Advances in School Mental Health Promotion*, *6*, 189–202.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Angold, A., Costello, E. J., Farmer, E., Burns, B., & Erkanli, A. (1999). Impaired but undiagnosed. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 129–137.
- Arnold, M. L., Newman, J. H., Gaddy, B. B., & Dean, C. B. (2005). A look at the condition of rural education research: Setting a direction for future research. *Journal of Research in Rural Education*, *20*, 1–25.
- Astramovich, R. L., Coker, J. K., & Hoskins, W. J. (2005). Training school counselors in program evaluation. *Professional School Counseling*, *9*, 49–54.
- Atkins, M. S., Fraizer, S. L., Adil, J. A., & Talbott, E. (2003). School-based mental health services in urban communities. In M. D. Weist, S. W. Evans, & N. A. Lever (Eds.), *Handbook of school mental health: Advancing practice and research* (pp. 165–178). New York: Kluwer Academic/Plenum Publishers.
- Berry, J. O., & Jones, W. H. (1995). The parental stress scale: Initial psychometric evidence. *Journal of Social and Personal Relationships*, *12*, 463–472.
- Bird, H. R., Shaffer, D., Fisher, P., & Gould, M. S. (1993). The Columbia Impairment Scale (CIS): Pilot findings on a measure of global impairment for children and adolescents. *International Journal of Methods in Psychiatric Research*.
- Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. (1997). The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(4), 545–553.
- Brady, C. E., Evans, S. W., Berlin, K. S., Bunford, N., & Kern, L. (2012). Evaluating school impairment with adolescents using the classroom performance survey. *School Psychology Review*, *41*(4), 429–446.
- Centers for Disease Control and Prevention. (2013). Mental health surveillance among children—United States, 2005–2011. *Morbidity and Mortality Weekly Report*, *62* (supplement 2). Retrieved December 15, 2013 from <http://www.cdc.gov/mmwr/pdf/other/su6202.pdf>.
- Chakraborty, K., Biswas, B., & Lewis, W. C. (2000). Economies of scale in public education: An econometric analysis. *Contemporary Economic Policy*, *18*, 238–247.
- Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., & Krull, J. L. (2010). Evaluation of the brief problem checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology*, *78*(4), 526.
- Cook, B. G., & Odom, S. L. (2013). Evidence-based practices and implementation science in special education. *Exceptional Children*, *79*, 135–144.
- de Anda, D. (2007). Intervention research and program evaluation in the school setting: Issues and alternative research designs. *Children and Schools*, *29*, 87–94.
- Deno, S., Mirkin, P., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children*, *49*, 36–45.
- DuPaul, G. J., Rapport, M. D., & Perriello, L. M. (1991). Teacher ratings of academic skills: The development of the Academic Performance Rating Scale. *School Psychology Review*, *20*, 284–300.

- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327–350.
- Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*, 1167–1187.
- Fabiano, G. A., Pelham, W. E., Jr., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., ... Burrows-MacLean, L. (2006). A practical measure of impairment: Psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child and Adolescent Psychology, 35*(3), 369–385.
- Farmer, E., Burns, B., Phillips, S., Angold, A., & Costello, E. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric Services, 54*, 60–66.
- Foa, E. B., Johnson, K. M., Feeny, N. C., & Treadwell, K. R. (2001). The child PTSD symptom scale: A preliminary examination of its psychometric properties. *Journal of Clinical Child Psychology, 30*(3), 376–384.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586.
- Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32*(1), 50–55.
- Hintze, J. M., & Matthews, W. J. (2004). The generalizability of systematic direct observations across time and setting: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review, 33*, 258–270.
- Hoagwood, K., & Erwin, H. D. (1997). Effectiveness of school-based mental health services for children: A 10-year research review. *Journal of Child and Family Studies, 6*, 435–451.
- Hoyt, D. R., Conger, R. D., Valde, J. G., & Weihs, K. (1997). Psychological distress and help seeking in rural America. *American Journal of Community Psychology, 25*, 449–470.
- Hunsley, J., & Mash, E. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology, 3*, 29–51.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.
- Jones, A. R., Cook, T. M., & Wang, J. L. (2011). Rural-urban differences in stigma against depression and agreement with health professionals about treatment. *Journal of Affective Disorders, 134*, 145–150.
- Kazdin, A. E., Rodgers, A., & Colbus, D. (1986). The hopelessness scale for children: Psychometric characteristics and concurrent validity. *Journal of Consulting and Clinical Psychology, 54*(2), 241.
- Kimber, B., Sandell, R., & Bremberg, S. (2008). Social and emotional training in Swedish schools for the promotion of mental health: An effectiveness study of 5 years of intervention. *Health Education Research, 23*, 931–940.
- Klein, D., Dougherty, L., & Olino, T. (2005). Toward guidelines for evidence-based assessment of depression for children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*(3), 412–432.
- Klein, D., Ouimette, P., Kelly, H., Ferro, T., & Riso, L. (1994). Test-retest reliability of team consensus best-estimate diagnosis of Axis I and II disorders in a family study. *American Journal of Psychiatry, 151*, 1043–1047.
- Knoff, H.M., & Batsche, G.M. (1994, October). Project ACHIEVE: A collaborative, school-based school reform process improving the academic and social progress of at-risk and underachieving students. *Paper presented at Safe schools, safe students: A collaborative approach to achieving safe, disciplined and drug-free schools conducive to learning conference, Washington, DC*. Retrieved from <http://files.eric.ed.gov/fulltext/ED383963.pdf>.
- Kratochwill, T. R., Hoagwood, K. E., Kazak, A. E., Weisz, J. R., Hood, K., Vargas, L. A., & Banez, G. A. (2012). Practice-based evidence for children and adolescents: Advancing the research agenda in schools. *School Psychology Review, 41*, 215–235.
- Langley, A. K., Nadeem, R., Kataoka, S. H., Stein, B. D., & Jaycox, L. H. (2010). Evidence-based mental health programs in schools: Barriers and facilitators of successful implementation. *School Mental Health, 2*(3), 105–113.
- Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: development of a general scale. *Evaluation and Program Planning, 2*(3), 197–207.
- Little, M., Murphy, J. M., Jellinek, M. S., Bishop, S. J., & Arnett, H. L. (1994). Screening 4- and 5-year-old children for psychosocial dysfunction: A preliminary study with the pediatric symptom checklist. *Journal of Developmental & Behavioral Pediatrics, 15*(3), 191–197.
- McLeod, B. D. (2011). Relation of the alliance with outcomes in youth psychotherapy: A meta-analysis. *Clinical Psychology Review, 31*, 603–616.
- McMahon, R., & Frick, P. (2005). Evidence-based assessment of conduct problems in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*(3), 477–505.
- Michael, K., Wandler, J., & Quick, A. (2010). Assessment, support, and counseling center. *North Carolina Medical Journal, 71*, 389–390.
- Michael, K. D., Albright, A., Jameson, J. P., Sale, R., Massey, C., Kirk, A., & Egan, T. (2013). Does cognitive behavioural therapy in the context of a rural school mental health programme have an impact on

- academic outcomes? *Advances in School Mental Health Promotion*, 6, 247–262.
- Owens, J. S., Andrews, N., Collins, J., Griffith, J. C., & Mahoney, M. (2011). Finding common ground: University research guided by community needs for elementary school-aged youth. In L. Harter, J. Hamel-Lambert, & J. Millesen (Eds.), *Participatory Partnerships for Social Action and Research* (pp. 49–71). Dubuque, IA: Kendall Hunt Publishers.
- Owens, J. S., Murphy, C. E., Richerson, L., Girio, E. L., & Himawan, M. L. (2008). Science to practice in underserved communities: The effectiveness of school mental health programming. *Journal of Clinical Child and Adolescent Psychology*, 37, 434–447.
- Pelham, W., Fabiano, G., & Massetti, G. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 34(3), 449–476.
- Pelham, W. E., Jr., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31(2), 210–218.
- Peltonen, K., Qouta, S., Sarraj, E. E., & Punamäki, R. L. (2012). Effectiveness of school-based intervention in enhancing mental health and social functioning among war-affected children. *Traumatology*, 18, 37–46.
- Poirier, J. M., & Osher, D. (2006). Understanding the new environment of public school funding: How student support services are funded. In C. Franklin, M. B. Harris, & P. Allen-Meares (Eds.), *The school services sourcebook: A guide for school based professionals* (pp. 1077–1091). New York, NY: Oxford University Press.
- Power, T. J., Dombrowski, S. C., Watkins, M. W., Mautone, J. A., & Eagle, J. W. (2007). Assessing children's homework performance: Development of multi-dimensional, multi-informant rating scales. *Journal of School Psychology*, 45(3), 333–348.
- Pritchett, R., Kemp, J., Wilson, P., Minnis, H., Bryce, G., & Gillberg, C. (2011). Quick, simple measures of family relationships for use in clinical practice and research. A systematic review. *Family Practice*, 28, 172–187.
- Ribbe, D. (1996). Psychometric review of traumatic event screening instrument for children (TESI-C). *Measurement of stress, trauma, and adaptation* (pp. 386–387).
- Riley-Tillman, T., Kalberer, S., & Chafouleas, S. (2005). Selecting the right tool for the job: A review of behavior monitoring tools used to assess student response-to-intervention. *The California School Psychologist*, 10, 81–91.
- Rimland, B., & Edelson, S. M. (1999). *Autism treatment evaluation checklist (ATEC)*. San Diego, CA: Autism Research Institute.
- Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review*, 3, 223–241.
- Scahill, L., Riddle, M. A., McSwiggin-Hardin, M., Ort, S. I., King, R. A., Goodman, W. K., ... Leckman, J. F. (1997). Children's Yale-Brown obsessive compulsive scale: reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 844–852.
- Schulte, A. C. (2010). Measurement in school consultation research. In W. Erchul & S. Sheridan (Eds.), *Handbook of research in school consultation* (pp. 33–61). New York: Routledge.
- Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research: Multidisciplinary perspectives on conceptualization, measurement, and enhancement of treatment integrity. *School Psychology Review*, 38(4), 460–475.
- Schultz, B. K., & Evans, S. W. (2012). Sources of bias in teacher ratings of adolescents with ADHD. *Journal of Educational and Developmental Psychology*, 2, 151–162.
- Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A children's global assessment scale. *Archives of General Psychiatry*, 40, 1228–1231.
- Silverman, W., & Ollendick, T. (2005). Evidence-based assessment of anxiety and its disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 34, 380–411.
- Vander-Stoep, A., Weiss, N. S., McKnight, B., Beresford, S. A., & Cohen, P. (2002). Which measure of psychiatric disorder—diagnosis, number of symptoms, or adaptive functioning—best predicts adverse young adult outcomes? *Journal of Epidemiology and Community Health*, 56, 56–65.
- Walker, H. M., & Gresham, F. M. (2013). The school-related behavior disorders field: A source of innovation and best practices for school personnel who serve students with emotional and behavioral disorders. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (vol. 7, 2nd ed., pp. 411–440). Hoboken, NJ: Wiley.
- Walker, H. M., Steiber, S., Ramsey, E., & O'Neill, R. (1993). Fifth grade school adjustment and later arrest rate: A longitudinal study of middle school antisocial boys. *Journal of Child and Family Studies*, 2, 295–315.
- Watabe, Y., Stewart, J. L., & Owens, J. S. (2013). Effectiveness and sustainability of school-based intervention for youth with or at-risk for ADHD. *School Mental Health*, 5, 83–95.
- Wei, Y., Szumilas, M., & Kutcher, S. (2010). Effectiveness on mental health of psychological debriefing for crisis intervention in schools. *Educational Psychology Review*, 22, 339–347.
- Weissman, M. M., Orvaschel, H., & Padian, N. (1980). Children's symptom and social functioning self-report scales comparison of mothers' and children's reports. *The Journal of Nervous and Mental Disease*, 168(12), 736–740.
- Weist, M. (2005). Fulfilling the promise of school-based mental health: Moving toward a public mental health

promotion approach. *Journal of Abnormal Child Psychology*, 33, 735–741.

- Weisz, J. R., Chorpita, B. F., Frye, A., Ng, M. Y., Lau, N., Bearman, S. K., ... Hoagwood, K. E. (2011). Youth Top Problems: using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. *Journal of Consulting and Clinical Psychology*, 79(3), 369.
- Weisz, J. R., Doss, A. J., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology*, 56, 337–363.
- Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (Eds.). (2010). *Handbook of practical program evaluation—Third edition*. San Francisco, CA: Wiley.
- Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods*, 6, 413–429.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. (2011). *The program evaluation standards: A guide for evaluators & evaluation users* (3rd ed.). Sage: Los Angeles, CA.
- Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: reliability, validity and sensitivity. *The British Journal of Psychiatry*, 133(5), 429–435.

Brandon K. Schultz, Ed.D., is an Assistant Professor in the pediatric school psychology program at East Carolina University. He worked for over a decade as a school consultant and clinical research scientist on a series of studies examining school-based interventions for adolescents with ADHD. This work culminated in the multisite clinical trial of the Challenging Horizons Program, which demonstrated the long-term efficacy of training interventions for ADHD and the importance of treatment dosage. His current research focuses on technologies to support school mental health dissemination and implementation in public schools.

Clifton Mixon, M.A., is a doctoral candidate in Clinical Psychology at Ohio University. He is currently completing his pre-doctoral internship in Pediatric Psychology at the Children's Hospital of Orange County (CHOC), and is scheduled to finish his doctorate in December 2017. His

clinical and research interests focus on increasing access to mental health care particularly for under-served communities. Much of his work has sought to accomplish this goal through integration of mental health professionals into educational and medical settings.

Anne Dawson, M.S., is a doctoral candidate in Clinical Psychology at Ohio University and is starting her pre-doctoral internship in pediatric psychology at the University of Louisville, School of Medicine. Anne's research interests include investigating the relationships between children and their significant contexts (families, teachers, peers, schools) and the way these relationships influence their development, specifically within the context of children with neurodevelopmental disorders.

Craig Spiel, Ph.D., received his B.S. from Brigham Young University and his M.S. and Ph.D. from Ohio University under the mentorship of Dr. Steven Evans. His research focus is on evidence-based services for youth with ADHD in schools. Craig completed his clinical internship with the University of Oklahoma Health Science Consortium in 2016 and his Postdoctoral Residency in the Family Mental Health Program at the Oklahoma City VA Health Care System in 2017. He is currently working for Intermountain Healthcare in Mental Health Integration in Salt Lake City, Utah.

Steven W. Evans, Ph.D., is a professor in the Department of Psychology at Ohio University and Co-Director of the Center for Intervention Research in Schools. His research interests are school-based treatment development and evaluation for adolescents with attention deficit hyperactivity disorder (ADHD) and related problems. He developed the Challenging Horizons Program which is a school-based treatment program for middle and high school youth with ADHD. Dr. Evans has maintained federal research funding from the National Institute of Mental Health and Institute for Educational Sciences to support this work for several years. Dr. Evans is the Editor-in-Chief of the journal *School Mental Health* and is on the editorial board of other related journals. He is a licensed clinical psychologist in Ohio and works with graduate students interested in clinical child psychology.