Assessing Effort and Malingering with the Asian Client

Olga Cirlugea

What Is Malingering and Effort?

Malingering is defined as "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" by the DSM-IV-TR and recommendations for suspecting the presence of malingering are made in situations in which two or more of the following are observed: 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 (American Psychiatric Association, 2000 p. 683). Rogers (1997) identified three models of malingering that map onto the APA's definition: pathogenic (underlying mental disorder), criminological (antisocial and oppositional motivation), and adaptational (cost–benefit analysis in response to adversarial circumstances). In order to study this construct, three primary approaches to assessing memory malingering have been developed: measures containing response style indices (e.g., MMPI-2), measures with a cutoff score indicative of memory malingering (e.g., Digit Span), and measures specifically designed for the assessment of memory malingering (e.g., Statement Validity Tests; Iverson & Binder, 2000). A number of problems in evaluating research on these measures with Asian populations have been raised: researchers often don't parse out the ethnic makeup of participants (e.g., Bowden, Shores, & Mathias, 2006), do not conduct analyses of ethno-racial factors, and fail to compare scores for different ethnic groups (e.g., O'Bryant, Engel, Kleiner, Vasterling, & Black, 2007; Temple, McBride, Horner, & Taylor, 2003).

Although the literature utilizes both "malingering" and "effort" as terms to describe a feigning of symptoms, in this chapter the term "malingering" is used as the majority of the assessment instruments discussed seek to determine whether symptoms are feigned and the term better captures the intentional production of symptoms that are motivated by an external incentive. By contrast, the term "effort" implies that motivation, which may arise from different sources (e.g., fatigue, boredom with the administration) plays a role in the examinee's performance and this doesn't adequately address the primary component of intentionality in malingering.

O. Cirlugea, B.A. (✉)
Department of Psychology, University of Nevada,
Reno MS 0296, Reno, NV 89557, USA
e-mail: olgacirlugea@gmail.com

Why Assessment Is Important?

Accurate assessment of malingering is very important due to (a) the potential iatrogenic effects of an intervention delivered to someone who doesn't need it or of failing to deliver an intervention to someone who does need it and (b) the financial costs associated with malingering, which include both payments for unnecessary medical services but also cost associated with fraud (Garriga, 2007). Accurate assessment of malingering can be especially complicated with working with minorities such as Asians because the research is either scarce or nonexistent and because differences in performance may due to cultural differences when a measure has been applied to groups for whom it hasn't been validated. In line with the APA guidelines on working with ethnic, linguistic, and culturally diverse populations, clinicians should be able to justify the use of any measure for the detection malingering as well as to discuss any limitations in interpretations of results due to cultural and linguistic factors.

There are three types of research methodology used to study malingering: known-groups comparison, differential prevalence design, and simulation design, each with methodological advantages and limitations (Rogers, 1997). While simulation designs in which participants are given scenarios (e.g. you're a failing student, about to be expelled, but if you can convince the school that you have a mental disorder you may stay at the school) then instructed to fake symptoms in order to put on a convincing presentation of mental disorders are used most often and have high internal validity (and therefore will be the primary types of studies reviewed in this chapter), they have been criticized for their lack of external validity. In addition to the evaluation of measures for validity and reliability, measures such as malingering ones that seek to differentiate between those who feign symptoms and those who don't also must be tested for their sensitivity (identifying malingerers correctly) and specificity (identifying non-malingerers correctly).

Aims of the Chapter

Thus, the aim of this chapter is to explore the most commonly used and/or researched measures that strive to assess for malingering specifically in the context of their use with Asians and to generate evidence-based recommendations for the assessment of malingering with the Asian client. In addition to reviewing the research on these measures in-text below, Table 11.1 offers a summary of each measure and it's availability in relevant languages. Findings indicate that, for the most part, researchers aiming to validate malingering assessments have not specifically evaluated their psychometric properties with Asian populations, although Asians have not been completely excluded from such research studies. Asian versions of several instruments have been developed in foreign countries (e.g., China and Korea) and those have been demonstrated to be effective in discriminating malingerers from non-malingerers, though it should be noted that research conducted in other countries may not generalize to Asian-American clients. Clearly, further research on the use of malingering assessment instruments with Asian clients is warranted for many of the instruments described in this chapter. However, there is little evidence that cultural differences play a significant role in the way in which clients respond on such measures in a manner that reduces their accuracy in detecting malingering and therefore most of the instruments discussed below are acceptable for use with Asian clients.

Test of Memory Malingering (TOMM: Tombaugh, 1996)

The TOMM is a 50-item recognition test developed to distinguish exaggerated or faked memory impairment. Clients are administered two learning trials and an optional retention trial during which 50 line drawings of common figures are presented and have to be correctly identified in recognition panels which are subsequently shown to the client. The retention trial is only

Table 11.1 Malingering and effort tests

Test name	Original language	Available in Asian languages?[a]	Research conducted with Asians	Are cultural modifications necessary?
Test of Memory Malingering	English	Since this test is essentially a nonverbal test it can be utilized with Asian clients	Multiple studies have supported the use of this measure with Asians	No
The Medical Symptom Validity Test (1) and Nonverbal Medical Symptom Validity Test	English	No (1); Since the NVMSVT is a nonverbal test it can be administered in any language	Yes, limited research has been conducted	No
Digit Memory Test	English	Since this is a test employing numbers it can be utilized with Asian clients	Two studies have evaluated the DMT specifically with Asians	No
Rey 15 Items	English	No (letters and Roman numerals utilized are drawn from Latin alphabet)	One study evaluated the Rey 15 Items specifically with Asians	Yes—in the case of elderly clients; language issues may exist as well
Digit Span	English	Since this is a test employing numbers it can be utilized with Asian clients	One study evaluated the Digit Span specifically with Asians	No
MMPI, MMPI-II, and MMPI-2-RF	English	Chinese, Korean, Hmong (MMPI-2 only)	Extensive research has been conducted with Asian Samples	Yes—please see the section on the MMPI and MMPI-2 for specifics
The Structured Interview of Reported Symptoms	English	No	Yes, limited research has been conducted	Yes—in the case of evaluation of adolescents
Miller-Forensic Assessment of Symptoms Test (M-FAST)	English	No	Yes, limited research has been conducted	No

[a]These include, but are not limited to, Chinese, Japanese, Korean, and those spoken by Southeast Americans (e.g., Hmong, Filipino, Thai, Burman, and Laotian)

administered if individuals score lower than 45 on the second learning trial. Individuals who receive a score lower than 45 on the second trial are considered to have an invalid test performance. Research reveals that over half (55.6 %) of participants in nonclinical populations score 49 or 50 on the TOMM and only 8.4 % score lower than 45 (Tombaugh, 1996).

The TOMM has been researched specifically with Asians in a single international study conducted in Hong Kong. Chang (2006) developed an assessment battery for the detection of malingering that included two measures developed by Chang herself (a famous faces test and

a subjective memory questionnaire), the TOMM, and indices in the Hong Kong List Learning Test (HKLLT, Chan & Kwok, 1999) indicative of memory malingering (i.e., Recognition Hits, Difference of Recall and Recognition, and False Alarm). Chang included the TOMM in her memory malingering battery of tests because the measure's stimuli "provided a cultural-free component" (p. 47). Two conditions were tested, simulated malingering (SM) and true effort (TE) in a sample of 57 community participants in Hong Kong. Significant differences between conditions were found for all four measures, and the battery assessment

correctly classified 96 % of participants in the SM condition and 100 % of those in the TE condition. The TOMM cutoff score of 45 correctly identified 84 % of participants in the SM condition and 100 % of those in the TE condition.

Subsequent data analyses compared the scores of 57 community participants to those of a sample of depressed patients ($n = 39$) in order to evaluate the effect of depression on memory malingering measures. Both groups were split into two the conditions described above, SM and TE; however, data from the simulated malingering Depression (SMD) group were dropped prior to analysis due to the failure of this group to comply with the feigning malingering instructions. Analysis of the remaining three groups, simulated malingering, true effort, and true effort depressed groups revealed significant differences between the SM and TE and SM and TED groups, and no significant differences between the two true effort groups on all measures of memory malingering except for on the HKLLT Recognition Hits index. The TOMM correctly identified 80 % of depressed patients and 100 % of normal controls.

A review of the extant literature on research of the TOMM in countries that have large immigrant populations (e.g., the USA, Canada) Asians have made up 0.008–40 % of the samples of studies where the TOMM was researched. For example, Moore and Donders (2004) examined the utility of the TOMM and the California Verbal Learning Test II (CLVT-II; Delis, Kramer, Kaplan, & Ober, 2000) in identifying invalid test performance in patients with Traumatic Brain Injury (TBI). The authors selected a sample of 132 patients, one of which was Asian-American who had suffered from a traumatic brain injury and had been seen at a Midwestern rehabilitation facility. Results indicated that the scores of 20 patients (15 %) were suggestive of invalid test performance (15 on the CLVT-II, 11 on the TOMM, and 6 on both measures). Further analysis of the data indicated that both financial compensation seeking and prior psychiatric history accounted for a large part of the invalid test performance.

Sollman, Ranseen, and Berry (2010) evaluated the ability of a battery of tests including ADHD self-report inventories, neuropsychological tests, Symptom Validity Tests (SVTs), and psychiatric malingering tests to detect feigning of ADHD symptoms in a sample 73 undergraduate students divided in three groups: an honest responding (clinically diagnosed) ADHD group (ADHD), a normal honest responding group (HON), and a normal feigning group (FGN). Three students, all in the FGN group, were Asian. SVT's included the TOMM, Digit Memory Test (DMT; Hiscock & Hiscock, 1989), Letter Memory Test, Card Version (LMT; Inman et al., 1998; Schipper, Berry, Coen, & Clark, 2008), and Nonverbal–Medical Symptom Validity Test (NV-MSVT; Green, 2006). Additionally, the Miller Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001) was administered to screen for psychiatric response rather than for malingering and "few participants endorsed many of the questions" on this measure (p. 331). Significant Differences were found between the ADHD and FGN groups on all four measures of malingering. The TOMM was determined to be the most promising in detecting malingering, with trial one revealing high sensitivity and moderate specificity, while the second learning trial and retention trial both revealed moderate sensitivity and high specificity. The remaining SVTs (DMT, LMT, and NV-MSVT) all had moderate sensitivity and high specificity. When results on multiple measures were combined and the threshold of three or more failures was selected, positive predictive power (PPP) of feigning was 100 %, highlighting the importance of using multiple measures in the assessment of malingering.

Gervais, Wygant, Sellbom, and Ben-Porath (2011) evaluated the relationship between Symptom Validity Test (SVT) failure and exaggeration of psychological symptoms on the MMPI-2 Restructured Form validity and substantitative scales (Higher Order [H-O], Restructured Clinical [RC], Specific Problems [SP], and Interest scales, and revised versions of the PSY-5 scales; MMPI-2-RF; Ben-Porath &

Tellegen, 2008). Of the 1,003 Forensic Disability Claimant participants, 2.7 % were Asian. Participants were administered several SVTs including the TOMM, the Word Memory Test (Green, 2003), the Computerized Assessment of Response Bias (Allen, Conder, Green, & Cox, 1997), and the Medical Symptom Validity Test (MSVT; Green, 2004). Only participants who had completed at least three SVTs were included for analysis ($n = 747$). Results indicated that a strong association existed between failure on the SVTs and overreporting of symptoms on the MMPY-2-RF as measured by the overreported psychological dysfunction (F-r), extreme psychopathology (Fp-r), somatic complaints (Fs), and noncredible reports of somatic or cognitive symptoms (FBS-r) scales.

Weinborn, Woods, Nulsen, and Leighton (2012) examined the effects of coaching on two Symptom Validity Tests, the MSVT (Green, 2004) and Nonverbal Medical Symptom Validity Test (NVMSVT; Green, 2007). Three conditions, symptom coached (SCG), test coaching (TCG), and combined symptom and test coaching (STCG), were compared to a control condition (best effort, BECG) in order to determine whether the original and nonverbal medical symptom validity tests were resistant to coaching interventions. These two STVs were then compared to a validated assessment instrument, the TOMM in a sample of 103 participants, which included 42 Asians (40.8 %). Specific to the TOMM, there was a statistically significant difference between SCG scores and TGC and STGC scores, with no significant difference emerging between the TGC and STGC conditions. Analyses of the MSVT found mixed results, with no statistical differences between TCG and SCG scores, statistically significant differences between STCG and TCG scores on four of the trials (Immediate Recall—IR, Delayed Recall—DR, Paired Associations—PA, and Free Recall—FR) and statistically significant differences between STCG and SCG scores on two of the trials (IR and FR). Lastly, results of the NVMSVT revealed no significant difference between TCG and SCG scores except for on the DRV and FR subtests, while there was a significant difference between STCG and SCG scores on IR, DR, Delayed Recognition-Associations (DRA), and Delayed-Recognition-Variations (DRV) trials but not on the Consistency (CNS) index. These findings indicate that the Best Effort Comparison Group's scores were "indicative of the highest levels of effort for all three SVT instruments, followed by the Symptom and Test Coached Group, then the Test Coached Group, with scores suggestive of the lowest levels of effort produced by the Symptom coached Group" (p. 839). Sensitivity scores were as follows: for the SCG, the TOMM and MSVT correctly identified 100 % of participants, while the NVMSVT correctly identified 96 %; for the TCG the TOMM 89 %, MSVT 92 %, NVMSVT 92 %; for the STCG the TOMM 83 %, 83 %, 88 %. These results indicate that both test and symptom coaching are required in order to produce better effort scores on the MSVT and NVMSVT, compared to just test coaching required for the TOMM.

With the exception of the international study conducted by Chang (2006), Asians have been largely excluded from research evaluating the TOMM. Additionally, given that the Chang (2006) study was conducted in Hong Kong, external validity issues may exist; the results may not generalize to Asians living in the USA; however, empirical evidence that suggests that the TOMM is not appropriate for use with Asians is absent. Therefore, the minimal research described above suggests that the TOMM is effective in detecting malingering and provides support for its use with Asian clients.

The MSVT (Green, 2004) and the NVMSVT (Green, 2007)

The MSVT is a shortened version of the Word Memory Test developed in order to assess effort. This measure contains three effort scales, Immediate Recall (IR), Delayed Recall (DR), and Consistency (CNS), and two memory scales, Paired Associates (PA) and Free Recall (FR). Effort is measured by the IR and DR scales by selecting between a correct and distracter word in

a forced choice task, while the CNS scales measures the degree of consistency between the IR and DR scores. A score of 85 % or less on any of the three scales is indicative of poor effort demonstrated by the patient. However, a significantly lower performance on the memory scales than on the effort scales is indicative of a dementia profile. The NVMSVT is a forced choice test containing of 10 paired objects (e.g., a cartoon drawings) presented on a computer screen for a total of 20 objects. Green (2007) reported 72.5 % sensitivity in the simulators, 95 % specificity in dementia patients, and 100 % specificity in volunteers providing good effort in the development of the measure.

The MSVT has been evaluated utilizing Asian clients as research participants and their inclusion in research studies has ranged from 2.7 to 40.8 %. Armistead-Jehle (2010) administered the MSVT to a sample of 45 veterans who had scored positive on a traumatic brain injury (TBI) measure at the Veterans Health Administration. Demographic analyses indicated that 6.7 % of the veteran sample was Asian, 51.1 % Pacific Islander, 22.2 % Caucasian, 15.6 % African and 4.4 % Hispanic. Results revealed that 58 % of participants received a failing score of 85 % or lower on at least one of the easy subtests (IR, DR, and CNS). Of importance, no significant differences emerged between those who passed and those who failed the MSVT among the various ethnic groups. Additional studies evaluating the MSVT include Gervais et al. (2011), and Weinborn et al. (2012) and discussed in the section on the "TOMM" above. The NVMSVT has not been studied specifically with Asians; however, the Weinborn et al. (2012) and Sollman et al. (2010) studies discussed above in the section on the "TOMM" have included Asians in their sample. The limited research on the MSVT and NVMSVT with Asians highlights the need for further studies in order better determine the instruments' psychometric properties with this population. However, the outcome of the Armistead-Jehle (2010) study lends support for the recommendation to use this measure when evaluating malingering with the Asian client.

Digit Memory Test (DMT; Hiscock & Hiscock, 1989)

The DMT utilizes a forced-choice paradigm in which individuals are first exposed to a card containing a number for 5 s, and then are asked to select that card from a two-card presentation containing the correct card as well as a foil. The procedure is repeated with 10 and 25 s presentation times. Performance on the DMT at less than chance levels (i.e. at the $p < 0.05$ level) is indicative of malingering.

Two international studies have evaluated the utility of the DMT with Asian clients. Chiu and Lee (2002) enlisted 38 Chinese university freshmen in order to determine whether detection of malingering behavior changed based on the level of task difficulty of the DMT. Easy items were ones in which the foil and correct response were produced by a random table of numbers (e.g., 42719 and 81359) while difficult items included numbers with just a few digits that were different (e.g., 62866 and 62686). All participants were tested in the control and fake bad conditions. Results revealed significant differences between the control and malingering group, the latter group performing more poorly than the former. Additionally, malingering participants had lower scores at the higher level of difficulty on the DMT. Cutoff scores were established based on Hiscock and colleagues' (1994) suggestion of actual group performance on the test. This resulted in a cutoff score of 29.7 that correctly classified 76 % of malingering participants at the difficult level and 30 % at the easy level. Based on these findings, the authors recommended that a more stringent cutoff score be used when the easy level of the DMT is tested while keeping in mind the probability of getting false-positives and false-negatives.

Liu, Gao, Li, and Sheng (2001) also found significant differences on the DMT between simulating malingerers and controls in a college sample, providing further support for the use of the instrument in detecting feigning of memory deficits. A third study, Sollman et al. (2010), described in the section on the "TOMM," included

A	B	C
1	2	3
a	b	c
○	□	△
I	II	III

Fig. 11.1 Rey 15 items, Benuto (2013)

Asians in the sample although scores were not reported by ethnicity. Given the scarcity of research on the psychometric properties of the DMT with Asian clients, further studies should be conducted to evaluate the ability of the DMT to detect malingering in Asians. Additionally, two of the three studies described were conducted in China, and there is always the possibility that the results may not generalize to Chinese living in the USA, other Asians, etc. However, there is currently no evidence that the DMT may function differently for various ethnic groups; therefore, its use is recommended with the Asian client.

Rey 15-Item Memory Test (Rey, 1964) and Wechsler Digit Span (Wechsler, 1990)

Rey 15 Item

The Rey 15 Item requires individuals to memorize 15 printed items (see Fig. 11.1; Benuto, 2013) after being exposed to them visually for 10 s, and to draw them on a blank piece of paper. The score is the number of correctly reproduced items. Salazar et al. (2007) has advocated for lowering cutoffs (recall plus [recognition minus false positives <20]) on the Rey 15 Item in order to avoid excessive false positives that may emerge due to differences between bilinguals and monolingual non-English-speaking patients. The Rey 15 has been evaluated in an international study (see Yamaguchi, 2005 below).

Wechsler Digit Span

The Wechsler Digit Span includes two tasks, digit forward and digit backwards. In the digit forward task, the administrator reads a series of digits (beginning with 2 and ending at 9) at the rate of 1 per second and the client is instructed to repeat those digits back to the administrator in correct order. After failure on two consecutive trials of a series, the test is discontinued. Researchers found that the ability to repeat numbers back in correct order is retained for elderly patients, those with dementia, and TBI (e.g., Benton, Esligner, & Damasio, 1981, Storandt, Botwinick, Danziger, Berg, & Hughes, 1984). In the digit backward task, the client is instructed to repeat the digits backward; the same discontinuation criterion is applied. A score of less than five on the digit span forward and less than four on the digit span backward is deemed atypical even for those elderly individuals aged 85–89 (Wechsler, 1997).

Both Digit Span and the Rey 15 item have been assessed with an Asian population. Specifically, Yamaguchi (2005) evaluated the Rey 15-Item Memory Test and the Wechsler Digit Span subset of the WAIS-R (Wechsler, 1990) in detecting malingered memory impairment in a Japanese sample. Fifty-two participants, including healthy young and elderly controls, instructed malingering, and elderly nursing home individuals were utilized for the study. Fifteen participants were assigned to Group I and received standard instructions, 17 participants were assigned to Group II and were instructed to lie about head-trauma symptoms resulting from a car accident, 12 elderly participants were assigned to Group III and received standard instructions as well and group IV was comprised of 8 nursing home participants, 7 with dementia and 1 with psychiatric disorder. The Rey 15-Item Memory Test instructions were translated into Japanese by the researcher, after which they were back translated into English by ten bilingual volunteers to ensure the accuracy of the instructions. The Wechsler Digit Span subset was taken from the Japanese version of the WAIS-R (Wechsler, 1990). After the administration participants in the malingering group were asked to determine if they thought they had successfully lied about having a head injury. Two separate analyses were performed with the first including all participants in the malingering group while the second excluding

participants who reported not being able to successfully malinger symptoms on both of the tests. Results for the Rey 15 Item revealed that the normal controls in group I performed similarly to the elderly in group III but significantly better than the malingerers in group II and the nursing home patients in group IV. Malingerers in group II and the elderly in group III both performed significantly better than the elderly in group IV. The most effective cutoffs were determined to be <9 for items and <2 for columns. Different sensitivity and specificity values were found for malingering based on which groups were used for comparison. Results of the Digit Span test indicated that normal controls in group I recalled significantly more items than malingerers in group II and nursing home elderly in group IV. There was no significant difference in number of items recalled between groups I and III, and groups II and IV. The most effective cutoffs were ≤8 digits for total digits, ≤4 and ≤5 digits for forward digit span, and ≤2 and ≤3 for backward digit span. Results were similar when the individuals who reported not successfully malingering were removed from the analyses. Effective cutoffs were ≤9 for items, and ≤1 and ≤2 for both columns and rows on the Rey 15 Item and ≤7 and ≤8 for total digit score, ≤4 for forward digit span and ≤2 and ≤3 for backward digit span. Taken together, these results provide most support for the utility of the Digit Span forward in identifying normal participants and discriminating them from the malingerers. The Rey 15 Item proved to be too difficult a task for the elderly nursing home residents with dementia. Because of this, the author recommended that the Digit Span be used when evaluating elderly impaired Japanese patients.

Neither the Rey 15 Item nor the Digit Span has been evaluated with Asians living in the USA. Therefore, the results of the Yamaguchi (2005) study may not generalize to those populations. However, there is no empirical evidence that ethnicity and culture significantly impact the ability of the Digit Span to detect malingering especially when items are drawn from a validated instrument (i.e., the Japanese version of the WAIS-R; Wechsler, 1990), and, therefore, its use

is recommended with Asian clients. In the case of the Rey 15 Item there is some concern that the items, which include symbols derived from the Latin alphabet, may not function in the way they were expected for those who are only familiar with Asian alphabets. Further studies should be conducted to determine whether the Latin symbols are adequate for use with Asian populations or whether the Rey 15 Item should be revised to include Asian symbols. The potential issue of language and Yamaguchi's (2005) finding that assessment instrument was too difficult for elderly nursing home residents with dementia indicate that caution should be used when the Ray 15-Item is employed with Asian clients; additionally, evaluation procedures should make use of multiple measures to assess malingering.

Minnesota Multiphasic Personality Inventory (MMPI) and Its Variants

The MMPI is one of the most frequently used measures in psychology to assess for personality and psychopathology. Because a large portion of the personality assessment chapter was dedicated to the MMPI-2 and the Asian client, this section will only focus on the MMPI-2 and the Asian clients as the measure is associated with malingering. Several MMPI scales have been used to detect malingering. The Fake Bad Scale (FBS; Lees-Haley, English, & Glenn, 1991) is the primary scale used for this purpose as it was specifically developed to assess response bias (Larrabee, 1998). Research has indicated that the FBS possesses good discriminant validity in studies of traumatic brain injury litigants (Greiffenstein, Baker, Gola, Donders, & Miller, 2002; Larrabee, 2003; Martens, Donders, & Millis, 2001; Ross, Millis, Krukowski, Putnam, & Adams, 2004). Additional studies reported in the Berry, Baer, and Harris (1991) and Rogers, Sewell, and Salekin (1994) meta-analyses have found that other scales such as the F (Infrequency) Scale, the F-K (Infrequency minus Defensiveness) index, the Fb (Back Infrequency Scale) Scale, the revised Social Desirability Scale, and the

obvious-subtle index were effective at detecting fake-bad profiles. Furthermore, the F and Fb scales and F-K index have been demonstrated to have good discriminate validity (Bagby, Nicholson, Buis, & Bacchiochi, 2000; Cramer, 1995; Lim & Butcher, 1996).

MMPI

With regard to the use of the MMPI and the Asian client, Sue and Sue (1974) compared the MMPI records of Chinese and Japanese students ($n = 48$) to those of non-Asian controls ($n = 120$) at a student clinic at a West Coast university and found that Asian students had elevated scores on the L and F scales. Results also revealed that the Asian students tended to underutilize mental health services, report a higher number of somatic complaints and to have higher familial discord. The authors hypothesized that the increased somatic complaints may be due to the acceptability of physical conditions over that of expression of mental problems in Asian families. Tsushima and Onorato (1982) conducted a similar study in which the medical records of white and Japanese-American patients at a private medical center in Hawaii were compared. Results indicated that there were no racial differences in scores on the different scales, and that other factors such as gender better accounted for any differences in responding. The authors noted that these results "imply that the MMPI interpretation rules based on white norms are applicable to certain Japanese-American medical patients" (p. 151) but that further research is needed in order for the results to generalize to other Asian-American population.

Wetter and colleagues (1992) attempted to determine whether validity scales on the MMPI-2 would be just as effective as the scales on its MMPI predecessor in differentiating random responses from malingering. In order to study the MMPI-2 scales, 151 college students, 7 % of which were Asian, were divided in four conditions: random responding, fake "moderate" group, fake "severe" group and control. Results indicated that the MMPI-2 scales were effective in differentiating between the two types of responding, especially the VRIN scale. Because the MMPI is in its second rendition, the research

discussed above was conducted 20+ years ago below more recent research on the MMPI-2 is discussed.

MMPI-2

Researchers found that Asians tend to endorse more somatic symptoms than Caucasians on the MMPI-2, for example in the case of depression (e.g., Marsella, Kinzie, & Gordon, 1971, Ryder et al., 2008). Tsushima and Tsushima (2002) evaluated whether racial differences existed between the MMPI-2 scores of 130 white and 66 Japanese-American outpatients at a private medical center in Hawaii. The authors found no significant differences on any of the 13 scales. However, power for the study was low (0.50) and therefore these results can't be interpreted to mean that Japanese-Americans score similarly to whites on the MMPI-2. A second study compared the scores of the outpatient sample to that of 32 "normals" that included two Japanese Americans, three Chinese Americans, one Korean American, and one Pacific Islander American. Specific to the validity scales, significant differences were found between the "normal" and outpatient sample, the latter scoring higher on scales F, K, and L. With regards to race, there were significant differences between Japanese-American Participants and "normals" on all scales with the exception of scale L and 5. Sue and colleagues (1996) evaluated differences in responding on the MMPI-2 among less acculturated Asians, highly acculturated Asians and Whites and the sample was 59.4 % Asian. The authors found no significant differences among the three groups on the Lie (L) and Defensiveness (K) scales, although there were significant gender differences on the L scale with females scoring higher than males. Differences were also significant on the Infrequency (F) scale, as less acculturated Asians scored significantly higher than their White counterparts. A pattern emerged among the three groups regardless of whether differences were significant or not, namely that the less acculturated Asians scored higher than the highly acculturated Asians, which in turn scored higher than the whites. This pattern was also observed in the profile validity of the three

groups. In light of these findings, the authors suggest that cultural factors be taken into consideration when interpreting the profiles of Asian-American clients, especially when the level of acculturation is low.

Tsushima and Tsushima (2009) sought to determine whether differences existed between Asian-American and Caucasian patients seeking compensation or participating in personal injury litigation on the MMPI-2 validity scales. Scores of 48 Asian-American medical patients were compared to those of 109 Caucasian patients on the following scales: F scale, Back Infrequency Scale, Symptom Validity Scale, Infrequency-Psychopathology Scale, and Dissimulation Scale. Results revealed no significant differences between the two ethnic groups on performance on any of the five scales.

Barber-Rioja, Zottoli, Kucharski, and Duncan (2009) evaluated the utility of the newly developed Criminal Offender Infrequency scale (Fc scale; Megargee, 2004) derived from the MMPI-2 in detecting malingering in a forensic sample, of which less than 1 % was Asian. The Structured Interview of Reported Symptoms (SIRS; Rogers, 1986) was used to classify 140 male criminal defendants as either malingering (23 %) or honest responders (77 %) using the malingering criteria of "three or more SIRS scales in the probable range or one scale in the definite range" (p. 19). Results demonstrated that the F, Fc, Fb, and F(p) scales all had acceptable sensitivity and specificity, with the Fc scale having similar sensitivity and specificity as the F and Fb scales. These findings provide support for the predictive utility of the newly developed Fc scale in detecting malingering.

MMPI-2-RF

The MMPI-2-RF (Ben-Porath & Tellegen, 2008) was developed in order to provide a more current evaluation of the models of psychopathology and personality that are presently being used. Gervais, Ben-Porath, Wygant, and Sellbom (2010) evaluated the incremental validity of the MMPI-2-RF symptom overreporting scales (F, Back Infrequency (Fb), Infrequency-Psychopathology (Fp), and FBS scales) and the scale developed

for the prediction of cognitive Symptom Validity Tests (SVT) scale, the Response Bias Scale (RBS; Gervais, Ben-Porath, Wygant, & Green, 2007) over the MMPI-2 in a sample of 1,187 non-TBI disability-related referrals and 2.7 % of them were Asian. Result indicated that the MMPI-2-RF scales had greater sensitivity in detecting memory complaints than their MMPI-2 counterparts. Building on the results of a previous study that demonstrated that elevated scores on the RBS were not associated with objective memory functioning (Gervais, Ben-Porath, Wygant, & Green, 2008), the authors point out that "subjective memory complaints in the context of elevated scores on RBS or the other MMPI-2-RF over-reporting scales are unlikely to indicate objective memory deficits, but rather suggest exaggerated memory or other cognitive complaints" (p. 281).

The Chinese MMPI

The Minnesota Multiphasic Personality Inventory (MMPI) was translated into Chinese in Hong Kong and China and the reliability of this version and its translation equivalence to the English version has been demonstrated in multiple studies (e.g., Boey, 1985; Cheung, 1985; National MMPI Coordinating Group, 1985). However, results have indicated that a number of scales including the F scale are elevated in Chinese populations (Cheung & Song, 1989). As a result, Cheung, Song, and Butcher (1991) developed two infrequency scales for the Chinese MMPI (ICH1 and ICH2) in response to the literature showing that Chinese American college students (Sue & Sue, 1974) and Chinese "normals" and psychiatric patients (Cheung, 1985, 1986; Cheung & Song, 1989; Cheung, Song, & Zhang, 1996; National MMPI Coordinating Group, 1982) all received elevated scores on the F scale on the original MMPI. Validation of the scales was conducted with a large sample including participants from Hong Kong (psychiatric patients, prisoners, and college students, some of whom were asked to fake good and some who were asked to fake bad) and China (convicted murderers, obsessive-compulsive, manic-depressive, and schizophrenic patients). Results indicated that the fake-bad group scored

extremely high on the Chinese infrequency scales, and significant differences between that group and the patient and prisoner groups emerged. The fake group's scores were similar to those of the normative sample. Among the two infrequency scales, the ICH1, which included items that were endorsed by no more than 10 % of Chinese and Hong Kong samples, emerged as better discriminating between normals and patients within a valid range. The authors recommend the use of this scale with Chinese patients over the original F scale even when this scale is rescored using local norms given that such transformations ignore the norms on which the scale was originally developed and complicate interpretations of score elevations.

The Korean MMPI-2 and Korean MMPI-2-RF

The MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) was first translated into Korean by Han (1993, 1996) and subsequently underwent multiple translations and back-translations until items were deemed adequate for testing with a Korean population. Hahn (2005) explored the utility of the validation scales of the Korean MMPI-2 (Kim et al., 2005) in the detection of fake-bad and fake-good profiles. The F (infrequency), Fb (fake-bad), F–K, and F(p) scales were developed when the original US MMPI scales resulted in elevated scores (e.g., high scores on F scale reported by Cho, Park, Ahn, & Shin, 1990). A sample comprised of South Korean students and psychiatric patients ($n=219$) was utilized for the study. The students were assigned to one of three conditions (fake-bad, fake-good problem denial, and fake-good claiming extreme virtue). Participants in the fake-bad condition received higher scores on the F, Fb and F(p) scales, and lower scores on the K scale compared to the psychiatric sample. The scales correctly classified between 87 and 95 % of faking-bad and psychiatric participants, with the F scale being most useful in this classification. On the other hand, faking-good profiles were classified with less accuracy, and of the scales, the S scale was the most useful in the classification. Because of the differences in the endorsement

of certain items between Americans and Koreans (e.g., #494—whether people should keep personal problems to themselves; #115—belief in after life; #20—satisfaction with sex life), the author recommended that instruments sensitive to a particular cultural context such as the Korean version of the MMPI-2 should be used. Dykhouse (2012) found that the Korean MMPI-2-RF validity scales developed from the Korean MMPI-2 item-pool correctly identified over or under-reporting honest participants in two conditions, uncoached and coached over (under)-report. Taken together, these results support the use of the Korean versions of the MMPI-2 and MMPI-2-RF. Further research is needed to determine whether these versions would be adequate for use with Koreans in the USA, particularly those who don't speak English or have a low level of acculturation.

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

The SIRS is the most widely used malingering measure in forensic practice (Archer, Buffington-Vollum, Stredny, & Handel, 2006). The second edition of the measure (SIRS-2; Rogers, Sewell, & Gillard, 2010) has been modified to prevent false-positive and false-negative classification errors (Rubenzer, 2010).

The SIRS has not been evaluated specifically with Asians; however, Asians have been part of the validation samples and their inclusion has ranged from 1 to 16.6 %. For example, Rogers, Gillis, Dickens, and Bagby (1991) evaluated the psychometric properties of the Structured Interview of Reported Symptoms (Rogers, 1986) in two studies. The first study evaluated the utility of the SIRS in differentiating simulators of malingering from two control groups, community and outpatient, and 7.4 % of the sample was defined as "Oriental." The second study evaluated potential differences in responding of suspected malingerers and psychiatric inpatients, of which 5.9 % were oriental. Responses on the SIRS were compared to results on the MMPI validity indicators and M Test. The authors found

that in the first study the SIRS was effective in discriminating between the two groups with excellent interrater reliability, while the second study revealed that suspected malingerers only responded similar to psychiatric inpatients on four of the nine SIRS scales (DS, DA, OS, and SO). The findings of the first study provide support for the discriminant and concurrent validity of the SIRS. In another study, Rogers, Gillis, Bagby, and Monteiro (1991) evaluated the ability of the SIRS to discriminate between coached and uncoached simulators of malingering in a sample of university students, of which 16.6 % were Asian. Scores were compared to those of the psychiatric sample utilized in the Rogers, Gillis, Dickens, et al. (1991) study. Students in both conditions were given a scenario in which they were instructed to feign mental illness in order to avoid expulsion from the university for failing grades, with the coached group receiving additional information about feigning mental illness. Even though the coached simulators were successful in reducing their malingering scores, all but one of the SIRS scales effective in discriminating them from uncoached and psychiatric counterparts. The SIRS accurately identified 100 % of uncoached simulators and 96.7 % of coached simulators, providing further evidence for the discriminant validity of the measures.

Rogers, Hinds, and Sewell (1996) examined the validity of three assessment instruments for the detection of feigning, the SIRS, the Minnesota Mulitphasic Personality Inventory-Adolescent (MMPI-A; Butcher et al., 1992) and Screening Index of Malingered Symptoms (SIMS; Smith, 1992), for use with an adolescent population. Fifty-three adolescents participating in a residential treatment program were recruited for the study, and one of them was Asian. Two conditions were tested, honest and feigning, the latter requiring participants to simulate symptoms of one of three disorders, schizophrenia, major depression, and generalized anxiety disorder. The authors found that they had to apply different rules for classifying malingering from the ones utilized in studies of MMPI and MMPI-2 with adult populations as these were inadequate when implemented to detect adolescent feigning.

Additionally, the SIRS demonstrated superior utility in classifying adolescents as either malingering or nonmalingering, with a smaller amount of false positives than the near one to one ratio found for the MMPI-A. These findings lent support for the clinical utility of the SIMS as a feigning screen for adolescents. Barber-Rioja et al. (2009) and Guy, Kwartner, and Miller (2006) have also utilized the measure in their studies, which can be found in the MMPI-2 section for the former and M-FAST section for the latter. Taken together, these results fail to reveal any evidence against using the SIRS with Asian clients; therefore, the use of the assessment instrument with this population is recommended.

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

The Miller-Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001) is a 25-item structured interview developed to detect malingering of psychotic symptoms. Scales include Unusual Hallucinations, Reported Versus Observed, Rare Combinations, Extreme Symptomatology, Negative Image, Unusual Symptom Course, and Suggestibility and the measure can be scored in one of three ways: the total score evaluates likelihood of malingering psychopathology; the scale scores inform clinicians about how the responders attempt to feign symptoms (i.e., though the reporting of unusual hallucinations); and several scales reliably discriminate malingerers from honest responders. The M-FAST has been shown to have good psychometric properties (Miller, 2001, 2004).

Similar to the SIRS, the M-FAST has not validated specifically with Asian clients. Nevertheless, some studies evaluating the psychometric properties of the M-FAST have included Asians in their sample. Guy et al. (2006) evaluated the psychometric properties of the M-FAST using a sample comprised of undergraduate simulators and psychiatric patients, the latter group containing 1.4 % Asians. Simulators were instructed to feign one of four mental disorders (schizophrenia, major

depressive disorder, bipolar disorder, and post-traumatic stress disorders), and malingering was identified using the SIRS and either the MMPI-2 or Personality Assessment Inventory (PAI; Morey, 1991). Results indicated that the M-FAST had excellent internal consistency but low mean inter-item correlation. Additionally, the measure was effective in discriminating simulators from their psychiatric counterparts, with the rare combination scale and the remaining scales varying in this ability, and feigned schizophrenia being more easily distinguished from true schizophrenia than the other disorders.

Messer and Fremouw (2008) investigated the sensitivity of the M-FAST and Morel Emotional Numbing Test-Revised (MENT-R; Morel, 1998) in 169 students. The sample was comprised of honest responders, coached malingerers (feigning PTSD after a car accident in order to obtain monetary compensation), and clinical PTSD responders, and 1.4 % of participants were Asian. Results indicated that the coached malingering group scored significantly higher on the MENT-R and the M-FAST than the other two groups. Additionally, the MENT-R accurately identified 63 % of malingers, while the M-FAST correctly identified 78 % of malingerers, with a combined correct identification rate of 90 %. The authors note that only 4 of 41 malingering participants met criteria for successful malingering, which may have influenced these results. While further research on psychometric properties of this measure with Asian samples is recommended, we currently have no evidence for why the M-FAST should not be part of the assessment of malingering with the Asian client.

Summary and Recommendations

In this chapter some of the most commonly used and well-researched tests that aim to assess for malingering and/or effort have been reviewed and a summary of these measures with appropriate recommendations can be found in Table 11.1. This review has revealed that there are several measures for malingering that have been specifically evaluated with Asian clients.

It is important to note that using a multi-method approach to the assessment of malingering as well as gathering collateral information will likely lead to increased accuracy of diagnoses that discern genuine psychological symptoms from symptoms of malingering. Additionally, limitations of the measures used, such as a lack of equivalency of a construct in the Asian client's native language, should be considered when administering a test that requires that results be interpreted with caution. In conclusion, clinicians should make use of the malingering assessment instruments described in this chapter in order to ensure that treatment is appropriately delivered and resources are not wasted on a client who does not in fact have a mental disorder.

References

Allen, L., Conder, R. L., Green, P., & Cox, D. R. (1997). *CARB '97 manual for the computerized assessment of response bias*. Durham, NC: CogniSyst.

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

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*, 85–94.

Armistead-Jehle, P. (2010). Symptom validity test performance in U.S. veterans referred for evaluation of mild TBI. *Applied Neuropsychology, 17*(1), 52–59.

Bagby, R. M., Nicholson, R. A., Buis, T., & Bacchiochi, J. R. (2000). Can the MMIP-2 validity scales detect depression feigned by experts? *Assessment, 7*(1), 55–62.

Barber-Rioja, V., Zottoli, T., Kucharski, L., & Duncan, S. (2009). The utility of the MMPI-2 Criminal Offender Infrequency (Fc) Scale in the detection of malingering in criminal defendants. *International Journal of Forensic Mental Health, 8*(1), 16–24.

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI–2–RF (Minnesota Multiphasic Personality Inventory–2–Restructured Form) manual for administration, scoring, and interpretation*. Minneapolis, MN: University of Minnesota Press.

Benton, A. L., Esligner, P. J., & Damasio, A. R. (1981). Normative observations on neuropsychological test performances in old age. *Journal of Clinical Neuropsychology, 3*, 33–42.

Benuto, L. (2013). Assessing effort & malingering with the Hispanic client. In L. Benuto (Ed.), *Guide to psychological assessment with Hispanics* (pp. 99–112). New York, NY: Springer.

Berry, D. T. R., Baer, R. A., & Harris, M. J. (1991). Detection of malingering on the MMPI: A meta-analysis. *Clinical Psychology Review, 11*(5), 585–598.

Boey, K. W. (1985). The MMPI response pattern of Singapore Chinese. *Acta Psychologica Sinica, 17*(4), 377–383.

Bowden, S. C., Shores, A. E., & Mathias, J. L. (2006). Does effort suppress cognition after traumatic brain injury: A re-examination of the evidence for the Word Memory Test. *The Clinical Neuropsychologist, 20*, 858–872.

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

Butcher, J. N., Williams, C. L., Grham, J. R., Archer, R., Tellegen, A. M., Ben-Porath, Y. S., & Kaemmer, B. (1992). *Minnesota Multiphasic Personality Inventory—Adolescent Version (MMPI-A)*. Minneapolis, MN: University of Minnesota Press.

Chan, A. S., & Kwok, I. (1999). *Hong Kong List Learning Test: Manual and preliminary norm*. Hong Kong: Chinese University of Hong Kong.

Chang, S. (2006). Development of a test battery for assessing memory malingering in Hong Kong and its application on depressed patients (China). *Dissertation Abstracts International, 68*.

Cheung, F. M. (1985). Cross-cultural considerations for the translation and adaptation of the Chinese MMPI in Hong Kong. In J. N. Butcher & C. D. Spielberger (Eds.), *Recent advances in personality assessment* (Vol. 8, pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cheung, F. M. (Ed.). (1986). *The Chinese Minnesota multiphasic personality inventory (MMPI): Research and applications* (Occasional Paper No. 12). Hong Kong: Centre for Hong Kong Studies, The Chinese University of Hong Kong.

Cheung, F. M., & Song, W. (1989). A review on the clinical applications of the Chinese MMPI. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*(3), 230–237. doi:10.1037/1040-3590. 1.3.230.

Cheung, F. M., Song, W., & Butcher, J. N. (1991). An infrequency scale for the Chinese MMPI. *Journal of Consulting and Clinical Psychology, 3*, 648–653.

Cheung, F., Song, W., & Zhang, J. (1996). The Chinese MMPI–2: Research and applications in Hong Kong and the People's Republic of China. In J. N. Butcher (Ed.), *International adaptations of the MMPI–2: Research and clinical applications* (pp. 137–161). Minneapolis, MN: University of Minnesota Press.

Chiu, V. Y., & Lee, T. C. (2002). Detection of malingering behavior at different levels of task difficulty in Hong Kong Chinese. *Rehabilitation Psychology, 47*(2), 194–203.

Cho, S., Park, B., Ahn, C., & Shin, D. (1990). The effectiveness of random response set identification scale of MMPI profiles. *Korean Journal of Clinical Psychology, 9*, 184–191.

Cramer, K. M. (1995). Comparing three new MMPI-2 randomness indices in a novel procedure for random profile derivation. *Journal of Personality Assessment, 63*(3), 514.

Delis, D. C., Kramer, J. H., Kaplan, E., Ober, B. A., et al. (2000). *California Verbal Learning Test—second edition*. San Antonio, TX: The Psychological Corporation.

Dykhouse, A. S. (2012). *Examining sensitivity of Korean MMPI-2-RF validity scales to over-reporting and under-reporting of symptoms*. Retrieved from http://condor.cmich.edu/cdm/singleitem/collection/p1610-01coll1/id/3553/rec/7

Garriga, M. (2007). Malingering in the clinical setting. *Psychiatric Times, 24*(3), 1–4.

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment, 14*, 196–208.

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2008). Differential sensitivity of the Response Bias Scale (RBS) and MMPI-2 validity scales to memory complaints. *The Clinical Neuropsychologist, 22*, 1061–1079.

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Sellbom, M. (2010). Incremental validity of the MMPI-2-RF over-reporting scales and RBS in assessing the veracity of memory complaints. *Archives of Clinical Neuropsychology, 25*(4), 274–284.

Gervais, R. O., Wygant, D. B., Sellbom, M., & Ben-Porath, Y. S. (2011). Associations between symptom validity test failure and scores on the MMPI-2-RF validity and substantive scales. *Journal of Personality Assessment, 93*(5), 508–517.

Green, P. (2003). *Green's Word Memory Test for Windows: Users manual*. Edmonton, AB: Author.

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

Green, P. (2006). *NV-MSVT for Windows, Version 1.3*. Edmonton, AB: Author.

Green, P. (2007). *Manual for the Nonverbal Medical Symptom Validity Test*. Edmonton, AB: Author.

Greiffenstein, M. F., Baker, W. J., Gola, T., Donders, J., & Miller, L. (2002). The fake bad scale in atypical and severe closed head injury litigants. *Journal of Clinical Psychology, 58*(12), 1591–1600.

Gualtieri, C., & Johnson, L. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology, 21*(7), 623–643.

Guy, L. S., Kwartner, P. P., & Miller, H. A. (2006). Investigating the M-FAST: Psychometric properties and utility to detect diagnostic specific malingering. *Behavioral Sciences and the Law, 24*(5), 687–702.

Hahn, J. (2005). Faking bad and faking good by college students on the Korean MMPI-2. *Journal of Personality Assessment, 85*, 65–73.

Han, K. (1993). The use of the MMPI–2 in Korea: Inventory adaptation, equivalence evaluation, and initial validation. *Dissertation Abstracts International, 54*.

Han, K. (1996). Adaptations in Asia and Southeast Asia. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 67–218). Minneapolis, MN: University of Minnesota Press.

Hiscock, C. K., Branham, J. D., & Hiscock, M. (1994). Detection of feigned cognitive impairment: The two-alternative forced choice method compared with selected conventional tests. *Journal of Psychopathology and Behavioural Assessment, 16*, 95–110.

Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology, 11*, 967–974.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10*, 128–149.

Iverson, G., & Binder, L. (2000). Detecting exaggeration and malingering in neuropsychological assessment. *Journal of Head Trauma Rehabilitation, 15*, 829–858.

Kim, I. J., Han, K., Lim, J., Lee, J., Min, B., & Moon, K. (2005). *Korean MMPI-2 user manual*. Seoul, Korea: Maumsarang.

Larrabee, G. J. (1998). Somatic malingering on the MMPI and MMPI-2 in personal injury litigants. *The Clinical Neuropsychologist, 12*(2), 179–188.

Larrabee, G. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17*(3), 410–425.

Lees-Haley, P. R., English, L. T., & Glenn, V. J. (1991). A fake bad scale on the MMPI-2 for personal injury claimants. *Psychological Reports, 68*, 203–210.

Lim, J., & Butcher, J. N. (1996). Detection of faking on the MMPI-2: Differentiation among faking-bad, denial, and claiming extreme virtue. *Journal of Personality Assessment, 67*(1), 1–25.

Liu, R., Gao, B., Li, Y., & Sheng, L. (2001). Simulated malingering: A preliminary trial on Hiscock's Forced-Choice Digit Memory Test. *Chinese Journal of Clinical Psychology, 9*(3), 173–175.

Marsella, A. J., Kinzie, D., & Gordon, P. (1971, March). *Depressive patterns among American college students of Caucasian, Chinese, and Japanese ancestry*. Paper presented at the Conference on Culture and Mental Health in Asia and the Pacific, Honolulu, Hawaii.

Martens, M., Donders, J., & Millis, S. R. (2001). Evaluation of invalid response sets after *traumatic* head injury. *Journal of Forensic Neuropsychology, 2*(1), 1–18.

Megargee, E. I. (2004, May). *Development and initial validation of an MMPI-2 infrequency scale (Fc) for use with criminal offenders*. Paper presented at the 39th Annual Symposium on Recent Developments on the MMPI-2/MMPI-A, Minneapolis, MN.

Messer, J., & Fremouw, W. (2008). Detecting malingered Posttraumatic Stress Disorder using the Morel Emotional Numbing Test-Revised (MENT-R) and the Miller Forensic Assessment of Symptoms Test (M-FAST). *Journal of Forensic Psychology Practice, 7*(3), 33–57.

Miller, H. A. (2001). *M-FAST: Miller Forensic Assessment of Symptoms Test*. Lutz, FL: Psychological Assessment Resources.

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*, 268–280.

Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury, 18*(10), 975–984.

Morel, K. R. (1998). *Manual for the Morel Emotional Numbing Test for Posttraumatic Stress Disorder (MENT)*. Printed in USA.

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

National MMPI Coordinating Group. (1982). The revision, employment and evaluation of MMPI in China. *Acta Psychologica Sinica, 14*, 449–458.

National MMPI Coordinating Group. (1985). Analysis of results of MMPI on normal Chinese subjects. *Acta Psychologica Sinica, 17*, 346–355.

O'Bryant, S. E., Engel, L. R., Kleiner, J. S., Vasterling, J. J., & Black, F. W. (2007). Test of Memory Malingering (TOMM) Trial 1 as a screening measure for insufficient effort. *Archives of Clinical Neuropsychology, 21*, 511–521.

Rey, A. (1964). *L'examen Clinique en Psychologie*. Paris: Presses Universitaires de France.

Rogers, R. (1986). *Structured interview of reported symptoms (SIRS)*. Clarke Institute of Psychiatry: Toronto, Unpublished scale.

Rogers, R. (1992). *Structured interview of reported symptoms*. Odessa, FL: Psychological Assessment Resources.

Rogers, R. (1997). Current status of clinical methods. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 373–379). New York, NY: Guilford Press.

Rogers, R., Gillis, J., Bagby, R., & Monteiro, E. (1991). Detection of malingering on the Structured Interview of Reported Symptoms (SIRS): A study of coached and uncoached simulators. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(4), 673–677.

Rogers, R., Gillis, J., Dickens, S. E., & Bagby, R. (1991). Standardized assessment of malingering: Validation of the Structured Interview of Reported Symptoms. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(1), 89–96.

Rogers, R., Hinds, J. D., & Sewell, K. W. (1996). Feigning psychopathology among adolescent offenders: Validation of the SIRS, MMPI-A, and SIMS. *Journal of Personality Assessment, 67*(2), 244.

Rogers, R., Sewell, K. W., & Salekin, R. T. (1994). A meta-analysis of malingering on the MMPI-2. *Assessment, 1*(3), 227–237.

Rogers, R., Sewell, K. W., & Gillard, N. D. (2010). *Structured interview of reported symptoms (SIRS)* (2nd Ed., professional manual). Lutz, FL: Psychological Assessment Resources.

Ross, S. R., Millis, S. R., Krukowski, R. A., Putnam, S. H., & Adams, K. M. (2004). Detecting incomplete effort on the MMPI-2: An examination of the Fake-Bad Scale in mild head injury. *Journal of Clinical & Experimental Neuropsychology, 26*(1), 115–124 & 10.

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

Ryder, A. G., Jian, Y., Xiongzhao, Z., Shuqiao, Y., Jinyao, Y., Heine, S. J., & Bagby, R. (2008). The cultural shaping of depression: Somatic symptoms in China, psychological symptoms in North America? *Journal of Abnormal Psychology, 117*(2), 300–313.

Salazar, X. F., Lu, P. H., Wen, J., & Boone, K. B. (2007). The use of effort tests in ethnic minorities. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 405–427). New York, NY: Guilford Press.

Schipper, L. J., Berry, D. T. R., Coen, E., & Clark, J. A. (2008). Cross validation of a manual form of the Letter Memory Test using a known groups methodology. *The Clinical Neuropsychologist, 22*, 345–349.

Smith, G. P. (1992). *Detection of malingering: A validation study of the SLAM test.* Unpublished dissertation, University of Missouri—St. Louis.

Sollman, M. J., Ranseen, J. D., & Berry, D. R. (2010). Detection of feigned ADHD in college students. *Psychological Assessment, 22*(2), 325–335.

Storandt, M., Botwinick, J., Danziger, W., Berg, L., & Hughes, C. P. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer type. *Archives of Neurology, 41*, 497–499.

Sue, S., Keefe, K., Enomoto, K., Durvasula, R., & Chao, R. (1996). Asian American and White college students' performance on the MMPI-2. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 206–218). Minneapolis, MN: University of Minnesota Press.

Sue, S., & Sue, D. W. (1974). MMPI comparisons between Asian-American and Non-Asian students utilizing a student health psychiatric clinic. *Journal of Counseling Psychology, 21*(5), 423–427.

Temple, R. O., McBride, A. M., Horner, M. D., & Taylor, R. M. (2003). Personality characteristics of patients showing suboptimal cognitive effort. *The Clinical Neuropsychologist, 17*, 402–409.

Tombaugh, T. N. (1996). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*, 260–268.

Tsushima, W. T., & Onorato, V. A. (1982). Comparison of MMPI scores of White and Japanese-American medical patients. *Journal of Consulting and Clinical Psychology, 50*(1), 150–151.

Tsushima, W. T., & Tsushima, V. G. (2002). Comparison of MMPI-2 scores of white and Japanese American medical patients. *Advances in Medical Psychotherapy and Psychodiagnosis, 11*, 1–8.

Tsushima, W. T., & Tsushima, V. G. (2009). Comparison of MMPI-2 validity scales among compensation-seeking Caucasian and Asian American medical patients. *Assessment, 6*, 159–164.

Wechsler, D. (1990). *Japanese Wechsler Adult Intelligence Scale-Revised: Manual* (F. Shinangawa, S. Kobayashi, K. Fujita, & H. Maekawa, Trans.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III: Manual*. San Antonio, TX: Psychological Corporation.

Weinborn, M., Woods, S., Nulsen, C., & Leighton, A. (2012). The effects of coaching on the Verbal and Nonverbal Medical Symptom Validity Tests. *The Clinical Neuropsychologist, 26*(5), 832–849.

Wetter, M. W., Baer, R. A., Berry, D. R., Smith, G. T., & Larsen, L. H. (1992). Sensitivity of MMPI-2 validity scales to random responding and malingering. *Psychological Assessment, 4*(3), 369–374.

Yamaguchi, T. (2005). *Detecting malingered memory impairment using the Rey 15-Item Memory Test and the Wechsler Digit Span subtest in a Japanese population*. UMI Dissertation Services, ProQuest Information and Learning, Ann Arbor, MI.