Assessment Instruments Measuring Malingering Used With Individuals Who Have Mental Retardation: Potential Problems and Issues

Kolleen E. Hurley and William Paul Deal

Abstract
Malingering, the exaggeration or fabrication of physical and/or psychological symptoms, can threaten the psychological assessment process (American Psychiatric Association, 2000). To enhance the validity of psychological evaluations, researchers have studied trends in malingering and developed instruments for its detection (Rogers, Bagby, & Dickens, 1992; Tombaugh, 1996). These instruments, however, may not be appropriate for individuals with significant subaverage intellectual functioning. Four instruments assessing malingering, frequently used in forensic evaluations, were administered to individuals with mental retardation. Results show that by utilizing established cut-off scores, we were able to classify a significant percentage of participants as “malingering,” in spite of directions to perform optimally. Practical implications as well as directions for future research are discussed.

Across contexts, malingering has been regarded as behavior motivated by a desire to avoid legal punishment, work, or other unfavorable duties or to obtain monetary awards (Cunnien, 1997). In the Diagnostic and Statistical Manual of Mental Disorders (4th ed., text revision), malingering is defined as “the intentional production of false or grossly exaggerated physical or psychological symptoms” (American Psychiatric Association, 2000, p. 739) with external gain as motive. The prevalence of malingering appears to vary by context (Rogers, 1997a), ranging from 5.6% in posttraumatic stress disorder assessments (Resnick, 1997) to 15.7% in forensic settings (Clark, 1997). These rates are substantial and very likely underestimate the actual prevalence of malingering (Rogers, 1997a; Youngjohn, Burrow, & Erdal, 1995). Assessment for malingering is considered an essential element of a forensic evaluation (Rogers, 1997a).

There have been a number of methods and approaches devised for the assessment of malingering. It is common for evaluations to include the use of intellectual and neuropsychological measures (Pankratz & Binder, 1997). Increasingly, assessment methodology for evaluating malingering has included use of structured interviews (Rogers, 1997b). These varied approaches are discussed in more detail below.

One of the most popular paradigms for assessment of malingering of intellectual and neuropsychological abilities has been forced-choice testing (Pankratz & Binder, 1997), sometimes referred to as symptom validity testing (Pankratz, 1979). In symptom validity testing, an examinee’s performance using a two-choice option is compared to chance performance. If the subject’s performance is significantly below chance (50% correct), the examiner may conclude the examinee has intentionally chosen incorrect responses and may have been motivated to perform inadequately (Pankratz & Binder, 1997). In an attempt to improve cutting scores based on simple chance performance, the illusion of task difficulty has been added to increase test specificity (Hiscock & Hiscock, 1989; Pankratz & Binder, 1997). Utilizing empirical norms to aid in decision-making has produced instruments such as the Rey Memory for 15-Items Test—hereafter called the Rey Memory Test and the Dot Counting Test (Lezak, 1976), as well as the Test of Memory Malingering (Tombaugh, 1996).
The Rey Memory Test and the Dot Counting Test have both been developed as brief measures to detect persons who may have been feigning or malingering cognitive deficits (Lezak, 1976). The Rey Memory Test requires an individual to graphically reproduce 15 items from memory. Although the test instructions emphasize the complexity and difficulty of the task, the actual demands on the examinee are assumed to be minimal (Lezak, 1976). The Dot Counting Test requires an individual to count the number of dots presented on index cards. Again, the appearance of task difficulty exceeds the actual demands on the examinee. Research with these instruments has suggested that malingering may be an issue if a subject has accurately reported fewer than 9 of the 15 items on the Rey Memory Test or has exceeded the mean time it has taken for patients with brain damage to count the dots on the Dot Counting Test (Lezak, 1976; Paul, Franzen, Cohen, & Fremouw, 1992).

Developed and based on forced-choice testing, the Test of Memory Malingering employs a visual-recognition procedure, during which the assessor asks the participant to decide which of two pictures of common objects he or she already has seen (Tombaugh, 1996). Clinicians then compare the percentage of correct responses to norms of individuals who have been deemed cognitively intact and those who have been deemed cognitively impaired. A score significantly below that of the mean for individuals with cognitive impairment suggests the possibility of less than optimal effort and possibly malingering (Tombaugh, 1996).

In addition to intellectual and neuropsychological testing, practitioners have also utilized structured interviews, one alternative to the traditional interview, to detect malingering. An examiner using a traditional interview format may have often overlooked persons engaged in malingering. This may have been especially true in cases where the person has experienced actual distress (e.g., about the impending evaluation) that has been mistaken for genuine symptom presentation (Rogers, 1997b). The Structured Interview of Report Symptoms (Rogers et al., 1992) is, as the name suggests, a structured interview based on the paradigm of forced-choice testing. The authors have designed it to detect malingered or feigned mental illness in individuals by recording reports of rarely endorsed symptoms or patterns of symptoms (Rogers et al., 1992).

These instruments have shown varying degrees of success in the assessment of malingering in individuals with average intellectual functioning (Gothard, Viglione, Meloy, & Sherman, 1995). It has been unclear, however, whether their psychometric properties have extended to a population of individuals with below average intelligence or mild mental retardation. Our focus in the present study was to examine those properties.

In the United States, there are approximately 6.2 to 7.5 million individuals with mental retardation; between 26,500 and 32,500 in prison or residential facilities for offenders; and thousands more on probation, detained in local jails, or housed with individuals with mental illness (Davis, 2000). It has been estimated that between 2% and 10% of the jail population in the United States consists of persons with mental retardation (Davis, 2000). These individuals may experience more vulnerabilities when encountering the legal system (Goldman, 2001) and may attempt to conceal their level of functioning (Appelbaum, 1994), with potentially significant consequences. For example, consider an individual with below average intelligence or mild mental retardation not previously identified or diagnosed, who is charged with a capital offense and evaluated for competency and responsibility. As Rogers (1997a) has noted, some form of assessment for malingering is considered an essential element of a comprehensive forensic evaluation. An individual’s true level of competency or actual degree of responsibility may be negatively affected by test results, suggesting malingered effort. Further, suspicion of malingering may cause the examiner to overlook or minimize the possibility of a diagnosis of mental retardation. Accurate and valid assessment would be critical in a case such as this, particularly in light of Atkins v. Virginia (2002), which prohibits capital punishment for offenders with mental retardation.

Hayes, Hale, and Gouvier (1997, 1998) have addressed malingering in persons with mental retardation in two studies. In the first study, they administered the Rey Memory Test, Dot Counting Test, and the M Test (Beaber, Marston, Michelli, & Mills, 1985) to individuals with mental retardation from a state facility for the criminally insane (Hayes et al., 1997). The M Test consists of 33 true–false items designed to identify individuals who are malingering symptoms associated with schizophrenia (Beaber et al., 1985). Hayes and colleagues’ results suggest a need for new norms or possibly simplified tasks or instruction and led the researchers to con-
clude that “the three measures used in this study should not be used to detect malingering in defendants with mental retardation” (p. 576). In their second study, however, Hayes et al. (1998) included the Structured Interview of Support Symptoms and concluded, “the Structured Interview of Support Symptoms alone and combined with other common tests of malingering accurately discriminates [sic] malingerers in a mentally retarded criminal population and can validate clinical decision making” (p. 36).

In a related study, Bianchini, Mathias, and Greve (2001) were not able to demonstrate that lower cognitive functioning makes a subject more likely to “fail” a symptom validity testing. In fact, they were more likely to “pass” than were college students simulating cognitive impairment, indicating that those with less intellectual capacity may not be at a disadvantage when symptom validity testing is employed. Likewise, sophisticated subjects (i.e., persons of average or above average intellectual functioning) may have the ability to successfully malinger on tasks of which they possess theoretical knowledge, but have no such advantage on tasks with which they are unfamiliar and, therefore, may have been detected as malingerers (Franzen & Martin, 1996). These findings suggest that intelligence may not play as critical a role in the “ability” to malinger (Bianchini et al., 2001). In contrast, an older study by Goldberg and Miller (1986) found that persons with mental retardation had trouble meeting the cut-off score of 9 on the Rey Memory Test.

These results have been inconsistent and often contradictory. Further, there have been no psychometrically sound instruments developed specifically for use with people who have mental retardation. Malingering may manifest as feigned symptoms of a psychiatric disorder, memory impairment, or decreased cognitive functioning. Whatever the manifestation, the possibility of malingering may contribute to heightened skepticism regarding the results of the forensic evaluation (Rogers, 1997a). It is essential to determine whether common malingering assessment instruments are appropriate for use in a population of individuals with mental retardation. As part of a broader study, Hurley (2003) administered four commonly utilized tests to assess malingerers in a mentally retarded criminal population and can validate clinical decision making (p. 36).

### Method

#### Participants

The subject pool included 39 participants (25 males, 14 females), whose mean age was 44.9 years (range = 22 to 69). Approximately 96% were Caucasian; the remaining 4% were African American. Based on previous assessment results, 10 individuals (26%) had Full Scale IQs between 70 and 78; 20 individuals (51%), between 60 and 69; and 9 (23%), between 50 and 59. Participants resided in one of two residential facilities for individuals with mental retardation, both located in a rural Southeastern state. Based on case review and self-report, none of them had prior involvement with the criminal justice system. Participation was voluntary; no compensation was offered, nor were there consequences for not participating. Informed consent was obtained from the participant or guardian. In cases where a guardian provided informed consent, assent also was obtained from the participant.

We used four assessment instruments designed to detect malingering in a population of individuals with average intellectual functioning. These instruments were selected because they were the ones most frequently cited in the current research literature (Back, Boone, Edwards, Parks, Burgoyne, & Silver, 1996; Hayes et al., 1997, 1998; Rogers et al., 1992). A brief description of each assessment follows.

#### Instruments

**Structured Interview of Reported Symptoms.** This structured interview, consisting of 172 items, was developed by Rogers et al. (1992) as a means to aid in the detection of malingering in clinical and forensic settings. It has three components: detailed inquiries, repeated inquiries, and general inquiries. All components are contained within eight primary scales derived from the empirical literature. Detailed inquiries require the examiner to ask the individual about specific symptoms and their severity and are followed by repeated inquiries as a check for response consistency. General inquiries encompass a number of dimensions, such as general and specific problems, as well as symptom patterns.

The authors of the Structured Interview of Support Symptoms have emphasized a multimethod
Assessment of malingering

approach to assessing malingering. Consequently, they did not provide explicit rules for detecting malingering; results of the interview are the sole source of data. The developers did, however, offer two main criteria and a supplementary strategy for detection: (a) any scale exceeding the cutting score for the definite feigning range, (b) a score in the probable feigning range for three or more of the scales, or (c) a total score (all scales) exceeding the cutting score of 76 (Rogers et al., 1992). Meeting any of these criteria has warranted suspicion that the subject is malingering. The average reliability of individual Structured Interview of Support Symptoms scales has been reported to be .96 (Rogers, Gillis, Dickens, & Bagby, 1991).

Test of Memory Malingering. Tombaugh (1996) developed this test as an instrument to aid psychologists in differentiating persons who suffer genuine memory impairment from those exhibiting insufficient effort via a recognition test. The Test of Memory Malingering was developed for adults and contains 50 items with two learning trials and an optional retention trial. We used results from the two learning trials and not the optional retention trial, which has been deemed sufficient by the author of the Test of Memory Malingering.

During each of the two learning trials, the examinee is shown the 50 line drawings (referred to as target pictures), which are illustrations of common objects. Each picture is shown for 3 s, with a 1-s interval between presentation of each picture. After displaying all 50 target pictures, the examiner presents the examinee with 50 recognition panels, one at a time, and asks them to identify which of the two common objects on each panel has been present during the learning trial. A subject's score is determined by assigning one point for each correct answer, making the range of possible scores 0 to 50. Individuals exerting adequate effort will produce higher scores than individuals exerting insufficient effort or deliberately distorting their responses. According to the manual, a score below 45 on the second learning trial should be noted by the examiner as “probable” malingering (Tombaugh, 1996).

The Test of Memory Malingering has been valuable in that it has been both “sensitive to malingering [and] . . . insensitive to neurological impairments” (Tombaugh, 1996, p. 1). It has been shown to correctly classify 91% of all subjects in validation studies, and individuals have typically possessed the capacity to recognize common objects, such as those used in this test, regardless of mental deficiencies (Tombaugh, 1997). Consequently, empirical norms have been able to be utilized, an advantage over the typical criterion of below-chance performance.

Rey Fifteen-Item Memory Test. The Rey Memory Test is an instrument used to validate memory complaints through employment of a strategy similar to that of the Test of Memory Malingering in that the required task appears more difficult than the actual task demand (Lezak, 1976). The administration of this test begins with the examiner giving instructions stressing that the subjects will be asked to remember 15 different items, with both 15 and different emphasized. In actuality, there are only 3 or 4 concepts, however, that need to be remembered in order to accurately recall the 15 items.

Each item is presented to the subject for 10 s. Following a 10-s delay, the examiner asks the subject to reproduce the 15 items. Unless the subject is suffering from a seriously disabling condition, he or she should be able to recall 3 of the 5 lines or 9 of the 15 items (Lezak, 1995). A score below the cut-off of 9 items may be an indication of malingering (Lezak, 1976).

Rey Dot Counting Test. This test is used to assess individuals presenting with intellectual impairment or a specific visual–perceptual deficit (Lezak, 1976). The test is based on the theory that there is a positive correlation between number of failures and task difficulty. In other words, the subject should make more mistakes as task difficulty increases. The stimuli are presented in random order with respect to level of difficulty. An individual exerting normal effort should display a pattern of failures that varies with the corresponding degree of task difficulty.

The test consists of 6 index cards consecutively numbered; there are 7, 11, 15, 19, 23, and 27 dots on each card, respectively. The subject is shown each card and asked to count, as quickly as he or she can, the number of dots on the card. The cards are presented in the following order: 2, 4, 3, 5, 6, 1, and response times for each card are recorded. We utilized criteria suggested by Paul and colleagues (1992). Specifically, any subject exceeding 180 s for total decision time may be identified as deliberately not exerting sufficient effort on the particular task (Paul et al., 1992).

Procedure

Following informed consent and/or assent, participants were tested in a comfortable room at ei-
ther his or her residential facility or workplace. The
Structured Interview of Support Symptoms, Test of
Memory Malingering, Rey Memory Test, and Dot
Counting Test. Order of test administration was
counterbalanced and were administered to partici-
pants, who were allowed to take as many breaks as
necessary. For all participants, testing was complet-
eed in one session varying in length from 45 minutes
to 2 hours.

## Results and Discussion

Descriptive statistics for each of the measures
are presented in Table 1. The average score for this
sample on the Structured Interview of Support
Symptoms Total, the Test of Memory Malingering,
and the Rey Memory Test exceed the established
cut-off for malingering. The percentage of individ-
uals classified as malingering when using the estab-
lished cut-off is also presented. On the Structured
Interview of Support Symptoms, 53.8% (total
score), 30.8% (one definite scale), and 30.8%
(three probable scales) of the sample exceeded the
cut-off. Forty-one percent of the sample fell below
the cut-off (45) for malingering on the Test of
Memory Malingering, whereas nearly 89% (79.5)
fell below the cut-off (9) for malingering on the Rey
Memory Test. Finally, only one individual (2.6%)
exceeded the cut-off for malingering on the Dot
Counting Test. When individuals with IQ over 70
(n = 10) were excluded from the analyses, the per-
centage of individuals classified as malingering re-
mained comparable. On the Structured Interview
of Support Symptoms total, definite, and probable,
58.6%, 34.5%, and 31%, respectively, of the sample
exceeded the cut-off. Forty-five percent fell below
the Test of Memory Malingering cut-off (45),
whereas nearly 90 (89.7%) fell below the cut-off (9)
for malingering on the Rey Memory Test.

A more detailed analysis of these results may
prove informative. The Structured Interview of
Support Symptoms total score ranged from 2 to 238,
suggesting tremendous variability in responses.
When participant IQ was compared to Structured
Interview of Support Symptoms total score, how-
ever, there was almost no relationship, \( r = .03 \).
In other words, participant level of intellectual func-
tioning did not appear related to the ability to val-
didly respond to the Structured Interview of Support
Symptoms questions and prompts. One possible ex-
planation may be related to the structure and for-

t Table 1 Cut-Off Scores, Means, SDs, and Percentage
Classified as Malingering by Instrument

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Cut-off</th>
<th>Mean</th>
<th>SD</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIRSa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>&gt;76</td>
<td>93.8</td>
<td>61.9</td>
<td>53.8</td>
</tr>
<tr>
<td>Definite scales</td>
<td>≥1</td>
<td>.90</td>
<td>1.7</td>
<td>30.8</td>
</tr>
<tr>
<td>Probable scales</td>
<td>≥3</td>
<td>1.5</td>
<td>1.7</td>
<td>30.8</td>
</tr>
<tr>
<td>TOMMb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 2</td>
<td>&lt;45</td>
<td>43.6</td>
<td>8.4</td>
<td>41.0</td>
</tr>
<tr>
<td>REY MFITc</td>
<td>&lt;9</td>
<td>5.6</td>
<td>4.0</td>
<td>79.5</td>
</tr>
<tr>
<td>Dot Counting Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total errors per subject</td>
<td>—</td>
<td>3.1</td>
<td>1.9</td>
<td>—</td>
</tr>
<tr>
<td>Total time</td>
<td>≥180</td>
<td>64.8</td>
<td>32.6</td>
<td>2.6</td>
</tr>
</tbody>
</table>

aStructured Interview of Report Symptoms. bTest of
Memory Malingering. cRey Memory for 15-Items Test.

cut of this instrument, in which affirmative re-
ponses produce higher scores. Individuals with
mental retardation may be more inclined to a pos-
itive-response bias regardless of the question con-
tent (Sigelman, Winer, & Schoenrock, 1982). This
positive-response bias may result in significant neg-
ative consequences during interactions with the le-
gal system, for example, false confessions (Evering-
ton & Fulero, 1999; Perske, 2000). It is possible
that such a bias resulted in high Structured Inter-
view of Support Symptoms total scores.

On the Test of Memory Malingering, a sub-
stantial minority of participants (41%) fell below
the cut-off score of 45 for Trial 2. As with the
Structured Interview of Support Symptoms, there
did not appear to be a strong relationship between
participant IQ and Test of Memory Malingering Tri-
al 2 scores, \( r = -.03 \). This relationship, or lack
thereof, held when participant IQ was compared to
Test of Memory Malingering Trial 1 scores, \( r = -.09 \).
There likely are other factors that are relevant for
performance on the Test of Memory Malingering at
this level of intellectual functioning. As one re-
viewer noted, however, the Test of Memory Malin-
erg
Trial 1 to Trial 2, 6 of 39 participants demonstrated either no change or actual improvement. These results suggest improvement in the scores from Trial 1 to Trial 2, which may provide useful information about an examinee’s level of effort and motivation. These results also suggest the possibility of developing norms specific to individuals with below average intelligence that may more accurately reflect their level of ability and performance.

Results from administration of the Rey Memory Test were even more noteworthy. Nearly 80% of the sample scored below the cut-off criteria for malingering. The mean number of correct responses was 5.6, well below the recommended cut-off of 9 correct. For the entire sample, the relationship between participant IQ and Rey Memory Test scores approached statistical significance, \( r = .31, p = .054 \). When participants with IQ over 70 were removed, however, this relationship disappeared, \( r = .07 \), suggesting that other factors may be related to performance on the Rey Memory Test. In fact, the relationship between Test of Memory Malingering Trial 1 scores and Rey Memory Test scores was notable, \( r = .54, p < .001 \). As both tasks appear to assess somewhat similar abilities (i.e., short-term visual memory), these results are not surprising.

Of the four instruments utilized, only the Dot Counting Test produced results consistent with expectation. This test has been utilized primarily as a screening device for malingering of general cognitive impairment (Lezak, 1976). Only one participant exceeded the cut-off score of 180 s, suggesting that this test may have promise as a brief screening device for individuals of below average intelligence. We note that there are a number of possible strategies for scoring the data besides total time. One such strategy involves the use of an additional six cards with grouped dots and deriving a combination score of average time for ungrouped, average time for grouped, and total errors (Nelson et al., 2002). Further study of the Dot Counting Test could involve this approach with individuals who have mental retardation to determine whether its potential utility is maintained.

There are a number of limitations associated with this study that must be acknowledged. The small sample was a relatively restricted one (i.e., the participants were primarily Caucasian, lived in a residential facility, and had no prior criminal justice system involvement). It is possible that the pattern of results we found may not generalize to a broader population. Further research could involve the participation of individuals with below average intelligence who do not live in residential facilities or those with below average functioning who have experienced contact with the criminal justice system. It is possible that these individuals might produce responses more consistent with expectations.

Taken together, these data at a minimum suggest considerable caution be exercised when using these various instruments for assessment of individuals with below average intelligence or mental retardation. A substantial minority (and in some instances a majority) of participants would have been suspected of “malingering,” despite having been instructed to perform optimally. In a forensic context, the consequences of being labeled in such a fashion can be significant. True impairment may be overlooked or minimized, leading to greater vulnerability for the individual (Goldman, 2001).

Further research and/or assessment strategies are needed to determine whether these instruments can ever provide valid results for individuals with significantly subaverage intellectual functioning. Including these individuals in a normative sample to determine appropriate cut-off scores would be one strategy. The development of instruments with tasks appropriate to this population would be another. Until this happens, the use of these instruments with individuals functioning at the lower ranges of cognitive abilities will pose substantial problems for the legal system and its participants.

References
Bianchini, K. J., Mathias, C. W., & Greve, K. W.
Assessment of malingering

K. E. Hurley and W. P. Deal


Tombaugh, T. N. (1996). The Test of Memory Ma-
Assessment of malingering

K. E. Hurley and W. P. Deal


Received 2/15/05, first decision 7/11/05, accepted 6/26/05.