The Validity of the M Test in a Residential Forensic Facility

Gary C. Hankins, MD; George W. Barnard, MD; and Lynn Robbins

The validity of the M Test as a screening measure for malingering was assessed in a residental forensic treatment center. Clinically, malingering was recognized by the treatment staff as a significant problem among incompetent to proceed (ITP) defendants in this setting. A total of 79 ITP defendants completed the M Test under standard instructions. To assess the predictive validity of the M Test, its sensitivity and specificity were determined using three separate measures of malingering status. Using the original scoring procedure, the results provided estimates of sensitivity of 11, 31, and 29 percent and estimates of specificity of 67, 70, and 69 percent, respectively, for the three malingering status measures. A revised scoring procedure improved the sensitivity estimates to the statistically significant level of 86 percent for the malingering status measure that was defined as feigned or grossly exaggerated psychotic symptoms. The pattern of correlations between the M Test scales and a variety of other clinical measures suggested that, in this forensic treatment setting, the subject's pattern of responses to the M Test was primarily determined by the severity of cognitive impairment of that subject.

Past research on malingering has demonstrated the significance of dissimulation in clinical practice, particularly in the forensic setting.[1] The actual prevalence of malingering for forensic populations has been estimated from a low of 3.5 percent in the Clarke Institute forensic assessment unit in Toronto[2] to a high of 27.1 percent in a sample of 85 men referred for outpatient forensic evaluations.[3] Outside of the forensic setting, systematic studies to assess prevalence of malingering have been rare.[1] Gillis, Rogers, and Bagby[4] have suggested that this apparent lack of focus on detecting

malingering may be due to "... an overreliance on clinical methods that are highly dependent on honest disclosure by the patient." The development of reliable and valid assessment devices to measure malingering would be a significant contribution to the field.

In 1985 the M Test was first described as a simple test to assess for the malingering of schizophrenia.[5] The M Test consisted of three scales (C, S, and M scales) based upon 33 true/false items. The authors offered data documenting a 78.2 percent correct classification rate of simulating college students (n = 104) and an 87 percent correct classification rate for a sample of hospitalized male schizophrenic patients (n = 65). The M Test met with rapid and apparently un-

Dr. Hankins is affiliated with the Union Correctional Institution, Raiford, Florida. Dr. Barnard and Lynn Robbins are affiliated with the Department of Psychiatry, University of Florida. Reprint requests to Dr. Hankins, 3326 NW 5th Ave., Gainesville, FL 32607.

critical acceptance by some authorities as a valid test for use in a forensic setting.[6]

Significant questions about the validity of the M Test have subsequently been raised. Personal communication with the developer of the M Test[7] revealed that he no longer considered the first eight of the 33 questions (the confusion or C scale) to be a valid part of the test. In a study cited by Gillis, Rogers, and Bagby,[4] Smith[3] administered the M Test to 85 male prisoners referred for forensic evaluation. Using the recommended cutting score, the sensitivity and specificity of the M Test were estimated at 69.6 and 66.1 percent, respectively. This low level of efficiency in accurately classifying suspected malingerers and true patients led Smith to question the usefulness of the M Test as a possible screening test for malingering.

Gillis, Rogers, and Bagby[4] used both a simulation and a natural group (forensic sample) design to assess the predictive validity of the M Test. They found that the sensitivity of the M Test for the simulators was 79.8 percent while the sensitivity for suspected malingerers among actual forensic patients was 40.0

percent. They emphasized the danger of exclusive reliance on simulation studies for validating measures of malingering. They further concluded that the M Test in its present form is not a valid screening measure for assessing malingering. Recently, Rogers, Bagby, and Gillis[8] reported the results of work done on a new two optional scoring procedure that appeared to produce better results. Option A produced sensitivity and specificity estimates of 81.0 and 83.8 percent while scoring Option B produced estimates of 95.2 and 70.6 percent, respectively. The authors stated that with the new scoring procedures "... the M Test can provide a rapid and effective method of screening out most patients for whom there is very little likelihood of malingering" (p. 103). The implication appears to be that with the new scoring procedure the M Test is valid for general clinical use.

At least four methodological criticisms of this study can be raised based on information in the published reports.[2, 4, 8] First, the predictive validity of the M Test was assessed by comparing the performance of 25 suspected malingers gleaned from chart reviews of all defendants evaluated over a seven-year period with the performance of 26 consecutive nonmalingering inpatients from the same forensic unit. The M Test purports to be a screening test for malingered psychosis—not for general malingering. The authors do not clarify the scope of the definition of malingering employed in the selection of the subjects for their probable malingering group. What proportion of the suspected malingerers were malingering psychotic

Table 1
Sociodemographic Characteristics of Sample (N = 79)

Mean age	35 ± 10 years
Race (Black or Hispanic)	47 (59.5%)
Single, widowed, separated, divorced	75 (94.9%)
Education	
Less than 10th grade	19 (24.1%)
10th–11th grade	20 (25.3%)
High school graduate	26 (32.9%)
Some college	9 (11.4%)
College graduate	5 (6.4%)

Table 2
Age, Race, Education and WRAT Score by Malingering Status

Malingering Status Measure	Race (% Nonwhite)	Age	Education	WRAT
Forensic expert's global judgment				
Positive (n = 9)	44.4	34.0 ± 6.9	10.4 ± 2.8	40.6 ± 19.7
Negative (n = 70)	61.5	35.7 ± 10.7	11.2 ± 2.8	45.9 ± 19.4
Primary therapist's global judgment				
Positive (n = 13)	61.5	31.7 ± 9.8	9.8 ± 2.3	34.1 ± 15.7*
Negative (n = 66)	59.1	36.3 ± 10.3	11.3 ± 2.9	47.5 ± 19.3
Primary therapist's finding of malingering psychotic symptoms				
Positive (n = 7)	85.7	28.0 ± 7.2*	9.3 ± 2.2	31.9 ± 15.5
Negative (n = 72)	56.9	36.3 ± 10.3	11.2 ± 2.8	46.6 ± 19.3

* <.05.

Table 3
Malingering Status and M Test Scales

M Test Based Judgment	Global Malingering— Forensic Specialist		Global Malingering— Primary Therapists		Malingering of Psychotic Symptoms	
CUTMSCALE	Positive	Negative	Positive	Negative	Positive	Negative
Positive	1	23	4	20	2	22
Negative	8	47	9	46	5	50
	Sensitivity 11		Sensitivity 31		Sensitivity 29	
	Specificity 67		Specificity 70		Specificity 69	
	Phi .15		Phi .004		Phi −.01	
Option A	Positive	Negative	Positive	Negative	Positive	Negative
Positive	5	28	8	24	6	26
Negative	4	42	5	42	1	46
	Sensitivity .56		Sensitivity .62		Sensitivity .86	
	Specificity .60		Specificity .64		Specificity .64	
	Phi −.10		Phi .19		Phi .29*	
Option B	Positive	Negative	Positive	Negative	Positive	Negative
Positive	6	35	9	32	6	34
Negative	3	35	4	34	1	38
	Sensitivity .67		Sensitivity .69		Sensitivity .86	
	Specificity .50		Specificity .52		Specificity .53	
	Phi −.11		Phi .15		Phi .22	

* p < .05 Fisher's exact test.

symptoms as opposed to memory deficits, intellectual impairment, behavioral symptoms, etc.? Including subjects who are not attempting to malinger psychotic symptoms in the suspected malingerer group would tend to bias against the validity of the M Test in this research design. This same criticism applies to the Smith study.[3]

A second criticism related to the impact of base rates on test performance. By comparing equal numbers of malingering and nonmalingering subjects, a sample is constructed with a base rate of 50 percent. This base rate maximizes the predictive power of the instrument. Cutting scores developed in such a manner are not likely to be valid in the lower base rate situations likely to be found in the forensic settings where the M Test

Table 4
Correlation of M Test Scale Scores with Symptom Denial and Symptom Exaggeration Scales†

M Test Scales	Memory Sx Scales		Intellectual Sx Scales		Psychotic Sx Scales		Behavior Sx	
	Deny	Exaggerate	Deny	Exaggerate	Deny	Exaggerate	Deny	Exaggerate
M Scale	−.031	−.129	−.191	−.280	.005	.055	.020	.022
S Scale	−.257	.031	−.160	.130	−.269*	.148	−.062	.160
Rule-In	−.083	.036	−.117	−.134	−.042	.190	−.034	.219

*$p \leq .05$.
† In Tables 4 through 7 the CI variable in the Rule-In scale was constructed on the basis of the primary therapist's global malingering judgment.

Table 5
M Test Scale Means by Global Symptom Presentation Style Over Treatment Period

M Test Scales	Mostly Exaggeration and/or Fabrication (n = 14)	Both (n = 37)	Mostly Minimization and/or Denial (n = 18)
M Scale	2.5	2.9	4.1
S Scale	5.2	4.9	4.2
Rule-In Scale	2.7	2.5	2.5

Differences not statistically significant.

Table 6
M Test Scale Means by Diagnosis

M Test Scales	Schizophrenia (n = 36)	All Other Diagnoses (n = 43)	t	p
M Scale	4.2 ± 4.6	2.4 ± 3.2	2.01	<.05
S Scale	4.9 ± 3.0	4.7 ± 2.9	.44	ns
Rule-In Scale	2.8 ± 2.9	1.8 ± 2.1	1.56	ns

would be employed as a screening instrument.

A third criticism is that it is unclear how appropriate the nonmalingering comparison group was. The prevalence of malingering in their sample is low (3.5%), suggesting that a narrow definition of malingering is being used and that their "probable malingering" group contained only the most extreme cases. Further, individuals were included in the comparison group only if there was no evidence of malingering. If malingering is a continuous phenomenon, the method described amounts to the com-

parison of cases only at the extreme ends of a continuum. This is a procedure that would tend to exaggerate differences and inflate correlations and could result in a strong bias in favor of the validity of the M Test. A more appropriate comparison group might have been all cases evaluated during the same time period in the same clinic who were not classified as malingerers.

The fourth criticism has to do with the methodology of validating test scoring procedures. The authors report that they calculated the positive predictive power and the negative predictive power

Table 7
Pearson Correlation of M Test Scales by Age, Education, WRAT Reading Level and CADCOMP Scales

Independent Variables	M Scale	S Scale	Rule-In
Age	−.14	−.09	−.15
Education	−.13	−.26*	−.17
WRAT score	−.25*	−.27*	−.34**
Adversarial process	.33**	.22*	.38***
Courtroom behavior	.41***	.39***	.39***
Active relationship with lawyer	−.14	−.02	−.13
Perception of relationship with lawyer	.04	.23*	.06
Psychotic features	.59***	.57***	.57***
Cognitive disorders	.37***	.19	.40***
Affective disturbance	.32**	.36***	.36***
Past psychopathology	.21	.55***	.23*
Self-directed aggression since arrest	.38***	.44***	.42***
Other-directed aggression since arrest	.38***	.36***	.33**
Criminal history	.05	.16	.04
Childhood/educational problem	.40***	.45***	.36**
History alcohol problems	.09	.33**	.07
History drug problems	.07	.33**	.08
Alcohol use day of crime	.03	.06	.01
Drug use day of crime	.20	.28*	.17
Crime awareness	.03	.04	.09
Psychotic symptoms day of crime	.20	.44***	.21

* $p < .05$.
** $p < .01$.
*** $p < .001$.

of all 33 items of the original M Test. Since few items were efficient in both directions, two separate scales were constructed: a Rule-Out scale consisting of the 10 items with the highest negative predictive power and a Rule-In scale consisting of the 10 items with the highest positive predictive power. Scales developed in this manner will be biased in the direction of the highest possible sensitivity and specificity for that research sample. Validation of such a scoring procedure would require testing on an entirely new sample of defendants. Such cross-validation research does not appear to have been undertaken by the authors.

The current research conducted in a state residential forensic treatment facil-ity attempts to address these apparent limitations of past research. First, in addition to the usual global measures of malingering, a measure of malingered psychosis was included in the analysis against which the validity of the M Test could be assessed. Second, the "suspected malingering" group was compared with all other ITP patients treated in the same clinical setting during the study period. Finally, going beyond the strategies employed in past research, the divergent validity of the M Test was explored by examining the relationship between M Test performance and other clinical measures of cognitive capacity, diagnosis, and psychopathology.

Methods

Subjects for this study were 79 residents/patients who had been ordered to

the North Florida Evaluation and Treatment Center (NFETC) as incompetent to proceed (ITP) from courts throughout the state. NFETC is a 210-bed forensic facility in Gainesville, Florida that is operated by the Florida Department of Health and Rehabilitative Services. The subjects were consecutive ITP defendants admitted during the first six months of 1991. The sociodemographic characteristics of these subjects are summarized in Table 1.

The primary therapist for each defendant was instructed to send each subject to the assessment laboratory for testing as soon after admission as feasible. A small number of subjects admitted during the study period could not be tested due to the severity of their psychopathology; less than five subjects refused to participate.

When the subject arrived in the laboratory, a technician explained the purpose of the evaluation, obtained informed consent, administered the Wide Range Achievement Test (WRAT) to determine reading level, and oriented the subject to the computer for the administration of the CADCOMP. The CADCOMP test consists of 272 questions answered primarily in a yes/no, true/false, or multiple choice format with answers entered on a simplified computer key board.[9, 10] A battery of four purported screening tests for malingering, including the M Test, was administered using standard instructions. At the recommendation of the test's developer[7] the first eight questions (the confusion or C scale) of the M Test were omitted. If the defendant was found to

be unable to read, the technician read all test questions aloud. The testing process typically took one to two hours to complete and was often spread over two or three consecutive days to accommodate the capacities of the subject.

Within one week of completing the testing, the subject underwent a detailed semi-structured psychiatric interview by one member of the research team (GWB) who is an experienced forensic specialist and who had no prior knowledge of the subject or the testing results. Immediately after the interview and based entirely on information obtained during the interview, the forensic specialist rendered a global judgment about the presence or absence of malingering.

Finally, about four months after the completion of all testing, the primary therapist for each subject, with no prior knowledge of the goals and objectives of the research or the results of the M Test, was asked to review the subject's chart and complete a structured instrument called the "Symptom Distortion Study." This instrument obtained information about the therapist's view of each subject's style of symptomatic behavior over the entire course of treatment in terms of a consistent pattern of denying symptoms, fabricating symptoms, both denying and fabricating symptoms, or neither denying or fabricating symptoms. These ratings were obtained for each of four categories of psychiatric symptoms, i.e., memory symptoms, intellectual symptoms, psychotic symptoms, and behavioral symptoms. In addition, the primary therapist was asked to render a judgment about the presence or absence

of malingering in relationship to each symptom category. This instrument provided two measures of malingering status; a global rating of malingering (probably malingering at least one symptom in any symptom category) and a rating of malingering limited to specifically psychotic symptoms (probably malingering at least one psychotic symptom). The DSM-III-R description of malingering was used as the standard throughout.

In the analysis of results, the three dependent variables were the M scale, the S scale (also called the Rule-Out scale in the revised scoring procedure), and the Rule-In scale. Two scoring procedures were used. The original recommended cutting score of four (4) or higher on the M scale was used to determine malingering (CUTMSCALE) by the first procedure.[5] The second procedure was that recommended by Gillis, Bagby, and Rogers.[4] Here malingering was determined either by a score of 4 or greater on their Rule-Out scale *and* a score of 2 or greater on the Rule-In scale (OPTION A) or by a score of 4 or greater on the Rule-Out scale *and* a score of 1 or greater on the Rule-In scale (OPTION B). Due to the fact that we did not become aware of the revised scoring procedure until after the completion of the testing of all of the subjects, one item on the Rule-In scale was not administered in the current study (the first question on the C scale of C1—"Grey is my favorite color"). To deal with this in the analysis, a constructed variable was used. The constructed variable was rated as follows: if a subject was judged to be a probable malingerer by the corre-

sponding independent variable for measuring malingering, the subject was considered to have answered C1 as "true," otherwise the subject was considered to have answered C1 as "false." This procedure biased the results in favor of greater specificity and sensitivity using the OPTION A and OPTION B scoring procedures.

The independent variables were the forensic specialist's global judgment of malingering, the primary therapist's global rating of malingering, and the primary therapist's rating of malingering of psychotic symptoms. Additional variables used were reading level derived from the WRAT, sociodemographic characteristics, primary DSM-III-R Axis I diagnosis obtained from the clinical record, and the 18 clinical scales of the CADCOMP test.

Results

The prevalence of malingering in the sample was as follows: global rating by the forensic expert identified nine of 79 subjects as malingering (11.4%), global rating by the primary therapist found 13 of 79 (16.5%), and rating of malingering of psychotic symptoms by the primary therapist identified seven of 79 (8.9%). On the global rating the forensic expert and the primary therapists agreed on 85 percent of the cases. The Kappa estimate of agreement on the global ratings was .80. Table 2 shows the distribution of subject's race, mean age, mean education, and mean WRAT score by malingering status for each of the three malingering measures. Of the 12 comparisons, two comparisons reached signifi-

cance using t tests. There appears to be a consistent trend for malingerers to be younger, less educated, and score lower on the WRAT.

The results for the validity assessment of the original scoring procedure of the M Test (CUTMSCALE) are shown in Table 3. None of the comparisons approached significance at the $p < .05$ level based on the Fisher exact test. The sensitivity estimates ranged from 11.1 to 30.8 percent and the specificity estimate ranged from 67.1 to 69.7 percent. Limiting the definition of malingering to the malingering of psychotic symptoms did not improve the efficiency of the original M Test scoring procedures.

Table 3 also contains the results for options A and B. These revised scoring procedures consistently improved the sensitivity of the M Test over the CUTMSCALE procedure, but they did so at the expense of specificity. Further, considering only the two global measures of malingering, none of the four resulting two-way tables reached statistical significance; the distributions could have been the product of chance. Restricting the definition of malingering to the malingering of psychotic symptoms improved sensitivity. The results for scoring Option A did reach significance at the .05 level by the Fisher exact test. However, the specificity remained low such that the proportion of false positives to true positives was 4.3:1 for Option A and 5.7:1 for Option B.

The logic of the M Test is based on the notion that malingerers will endorse a high number of items soliciting valid and common psychotic symptoms; i.e.,

they will exaggerate on the S scale (Rule-Out scale) and they will tend to endorse items soliciting rare or bizarre psychotic symptoms on the M scale or the Rule-In scale. Table 4 shows the relationship between primary therapist rated scales assessing persistent patterns of denial or fabrication of symptoms throughout each subject's entire course of treatment for each of the four symptom categories with the M scale, the S scale, and the Rule-In scale (note: the Rule Out scale is not shown since it is identical with the S scale). If the M Test scales are operating in a manner consistent with the logic of the test's construction, the M Test scales should be positively associated with the symptom fabrication scales and negatively associated with the symptom denial scales. The only significant relationship was a negative correlation between S scale scores and ratings on the denial scale for psychotic symptoms. Thus, subjects which the primary therapist saw as tending to minimize valid psychotic symptoms also tended to endorse significantly fewer of the 10 items on the S scale, which was intended to assess the prevalence of legitimate psychotic symptoms.

Continuing this line of analysis, Table 5 shows the results of testing for differences in mean scale scores as a function of the primary therapist's Global Symptom Presentation Style judgment. No significant differences were found. It may be noteworthy that the lowest mean M scale scores were for subjects judged as demonstrating "mostly exaggeration and/or fabrication" on the primary ther-

apist's Global Symptoms Presentation Style rating.

Divergent validity of the M Test was assessed by examining the relationship between the M scale and Rule-In scale scores and a variety of clinical measures with which they should presumably be unrelated. Table 6 shows the result for the relationship with discharge diagnosis. As will be noted, the mean M scale score was significantly associated with a diagnosis of schizophrenia. Schizophrenic subjects had significantly higher mean M scale scores than nonschizophrenic subjects. S scale and Rule-In scale scores were not significantly associated with a diagnosis of schizophrenia.

The pattern of correlations between the M Test scale scores and subject age, education, WRAT scores, and the 18 clinical scales of the CADCOMP are summarized in Table 7. M scale scores were significantly negatively correlated with reading level as measured by the WRAT score. M scale scores were moderately and positively correlated with impaired knowledge of the adversarial legal process and with current affective distress. M scale scores were strongly and positively correlated with lack of knowledge of appropriate courtroom behavior, current psychotic features, current cognitive disorders, past aggressive impulse control problems, and self-reported educational problems in childhood. Not surprisingly, the Rule-In scale, which shares 9 items in common with the M scale, shows a nearly identical pattern of correlations as the M scale. The S scale showed a somewhat similar pattern of correlations with some notable exceptions. The S scale showed a significant negative association with educational level. There was not a significant correlation with current cognitive disorders but a strongly positive correlation with a past history of serious psychopathology and psychotic symptoms on the day of the alleged crime. Using a stepwise regression model, four CADCOMP scales, psychotic features, aggression toward self, cognitive disturbance, and active relationship with lawyer, predicted 48.9 percent of the M scale variance. For the Rule-In scale, four CADCOMP scales, psychotic features, aggression toward self, cognitive disturbance, and active relationship with lawyer plus the WRAT scale predicted 52.2 percent of the variance in the scale scores. Finally, five CADCOMP scales, psychotic features, past psychopathology, aggression toward self, use of alcohol on day of the crime, and use of drugs on day of the crime as well as the WRAT scale predicted 58.5 percent of the S scale variance.

Discussion

The major result of the current investigation is that the M Test was found to be lacking in predictive validity as a measure of malingering status in the current setting when the broadest definition of malingering was used as the standard. This was true regardless of whether the original or the revised scoring procedure was employed and despite the fact that the use of the constructed C1 item on the Rule-In scale biased the results in favor of the validity of the revised scoring procedure. Restricting the definition

of malingering to psychotic symptoms improved the sensitivity of the revised scoring procedure but did so at the cost of a considerable reduction in specificity. Using this malingered psychosis measure as the standard, the revised scoring procedure, Option A, reached statistical significance. Thus, the revised scoring procedure showed some promise of salvaging the M Test as a meaningful screening test for malingered psychosis.

Performance on the M scale was found to be significantly related to a discharge diagnosis of schizophrenia. Also striking was the high level of correlation between performance on the M scale, the S scale and the Rule-In scale with a variety of clinical scales indicative of serious current psychopathology and current cognitive impairments. Considering the lack of precision in the measures used, the multiple regression analysis revealed that most of the variation in M Test scale performance could be accounted for by these measures of clinical status. This result calls into question a central aspect of the logic underlying the construction of the M Test, i.e., that true patients should endorse few of the rare or bizarre symptom items (the M scale and the Rule-In scale) and that performance on these scales should, therefore, be unrelated to the clinical status of true patients.

In retrospect this correlation with current cognitive status is not altogether surprising. The strategy underlying the design of the M Test is essentially cognitive in nature. The subject is asked a true/false item, the subject "thinks" about the item, the subject gives a re-

sponse, and the subject is assigned points depending upon the content of their responses. The process described above suggests that forensic subjects may "think" about test items differently from what researchers and simulators do and/or that their current cognitive status significantly impacts their ability to engage in this cognitive enterprise. The same confounding effect of cognitive status might apply to all malingering assessment devices that are based on the scoring of forensic subjects' responses to "trick" questions of the type used in the M Test.

Gillis, Rogers, and Bagby[4] have already demonstrated and emphasized the danger of using a simulation design to validate new tests of malingering. The current research emphasizes additional cautions that may be of importance in assessing the validity of measures of malingering. Among the most important are the manner in which malingering status is operationally defined, the issue of the composition of the comparison group, and the appropriateness of generalizing the results obtained in one forensic setting to other forensic settings. Further, the current results suggest that, in future research, attention will need to be paid to the potentially powerful confounding effects of the subjects' concurrent levels of cognitive impairment.

The current study is now the second published study[4] that has failed to validate the M Test and the original scoring procedure using the natural group design. This result raises further doubt about the validity of the original scoring procedure. As a cross-validation of the

revised scoring procedure, the current study fails to confirm validity when the basis of judgment is global malingering. The results do suggest that if the use of the M Test is restricted to screening subjects for the malingering of psychotic symptoms, a useful level of sensitivity may be possible. The results also suggest that with this restricted application, the specificity of the M Test is sufficiently low (.50–.53) that for every case correctly classified there will be about five false positive cases, when the base rate of malingered psychosis is around 10 percent. Since this low level of specificity is to a large extent a consequence of the impact of the current cognitive impairments on test performance, it follows that the specificity of the M Test could well be found to be higher in settings where defendants are on the average less cognitively impaired than is typical at NFETC.

The above results reveal some promise for the M Test under the revised scoring procedure as a screening test for malingered psychosis. Although the M Test is of considerable research interest, in the opinion of the authors there is as yet insufficient data to support recommending it for routine clinical use.

References

1. Rogers R (ed): Clinical Assessment of Malingering and Deception. New York, Guilford Press, 1988
2. Rogers R, Gillis JR, Dickens SE, Bagby RM: Standardized assessment of malingering: validation of the structured interview of reported symptoms. Psychological Assessment 3:89–96, 1991
3. Smith GP: Detection of malingering of schizophrenia in male prisoners. J Psychol Assess, submitted
4. Gillis JR, Rogers R, Bagby RM: Validity of the M Test: simulation-design and natural-group approaches. J Pers Assess 57:130–40, 1991
5. Beaber RJ, Marston A, Michelli J, Mills MJ: A brief test for measuring malingering in schizophrenic individuals. Am J Psychiatry 142:1478–81, 1985
6. Grisso T: Competency to Stand Trial Evaluations: A Manual for Practice. Sarasota, FL, Professional Resource Exchange, Inc., 1988
7. Beaber RJ: Personal communication, 1990
8. Rogers R, Bagby M, Gillis JR: Improvements in the M Test as a screening measure for malingering. Bull Am Acad Psychiatry Law 20:101–4, 1992
9. Barnard GW, Thompson JW, Freeman WC, et al: Competency to stand trial: description and initial evaluation of a new computer-assisted assessment tool (CADCOMP). Bull Am Acad Psychiatry Law 19:367–81, 1991
10. Barnard GW, Nicholson RA, Hankins GC, et al: Itemmetric and scale analysis of a new computer-assisted assessment instrument (CADCOMP). Behav Sci Law 10:419–35, 1992