



## The detection of malingering in memory performance: The sensitivity and specificity of four measures in a UK population

P. J. Kelly<sup>1</sup>, G. A. Baker<sup>1\*</sup>, M. D. van den Broek<sup>2</sup>, H. Jackson<sup>3</sup>  
and G. Humphries<sup>4</sup>

<sup>1</sup>Neurological Sciences, Clinical Science Centre for Research  
and Education Liverpool, UK

<sup>2</sup>The Wolfson Neurorehabilitation Centre & St George's Hospital London, UK

<sup>3</sup>TRU Rehabilitation Services, St Helens, UK

<sup>4</sup>Department of Clinical Psychology, University of Manchester, UK

**Objectives.** To investigate the validity of a clinical neuropsychological battery for the detection of malingering on tests of memory.

**Methods.** A simulated scenario design was developed to investigate the effectiveness of a battery of four neuropsychological tests in the detection of malingering; the Coin in the Hand Test (CIH), Autobiographical Memory Index (AMI), Rey 15-Item Test (RIT), and the Wechsler Mental Control Test (MCT). The performances of patients with an acquired brain injury ( $N = 40$ ) were compared with two groups of controls instructed either to simulate a head injury performance ( $N = 40$ ) or do their best ( $N = 40$ ).

**Results.** The CIH and MCT demonstrated good validity and displayed high sensitivity and specificity. The RIT and the AMI was relatively poor in distinguishing between simulators and patients.

**Conclusions.** The sensitivity and specificity of all four tests to the detection of malingering has been assessed. Two of the tests the CIH and MCT would be useful as a quick and accurate screening tool for detecting malingering.

Greater demands are being placed on the medical profession to provide an opinion on patient malingering, specifically in the context of a claim for compensation for personal injury. In the case of head injuries, neuropsychologists are increasingly asked to undertake extensive neuropsychological assessments to determine whether the patient maybe attempting to exaggerate or fake genuine neuropsychological deficits (Resnik, 1984). Malingering has been defined as the intentional production of false or

\* Correspondence should be addressed to Professor Gus A. Baker, University Department of Neurosciences, Clinical Sciences Centre for Research and Education, Lower Lane, Fazakerley, Liverpool L9 7LJ, UK (e-mail: g.baker@liv.ac.uk).

greatly exaggerated symptoms for the purpose of attaining some identifiable external reward (American Psychiatric Association). While neuropsychological malingering has been defined specifically as the wilful production of a poor performance, this is usually displayed by highly unusual test results (Franzen, Iverson, & Cracken, 1990). The reason for this under performance can be either financial or emotional gain, or both. In the field of psychiatry and psychology the *Diagnostic Statistical Manual (DSM - IV)* of the American Psychiatric Association specifies that a diagnosis of malingering should be considered when the client is seen for medico-legal purposes, discrepancies are found between the level of distress/disability reported by the client and that found by the clinician, and the client is uncooperative and has an antisocial personality.

Increasingly sophisticated methodologies and protocols have been developed and applied to detect feigned, memory impairment (Pankratz, 1983; Graf, Squire, & Mandler, 1984; Schacter, 1986; Brandt, 1988; Binder & Willis, 1991). However as Rogers (Rogers, Harrell, & Liff, 1993; Rogers, 1997) has pointed out, many neuropsychological tests were not developed for the purpose of detecting malingered performance. Clinicians may also be concerned about the time it takes to administer additional neuropsychological tests that have been developed to detect malingering. Moreover, the neuropsychologist is now presented with the difficulty of deciding which test to administer to obtain the best chance of determining whether or not an individual is malingering. A recent publication by Hutchinson (2001) has reviewed a number of pertinent issues relating to the assessment of malingering.

In this study the authors have drawn upon their previous work (Baker, Hanley, Jackson, Kimmance, & Slade, 1993; Cochrane, Baker, & Meudell, 1998; Hanley, Baker, & Ledson, 1999) to identify a selection of tests that could be used to determine whether an individual is malingering. The authors identified four tests commonly used in the UK that might have the potential for inclusion in a battery for assessing malingering. These are the Coin in the Hand Test (CIH; Kapur, 1994), the 11 Question Autobiographical Memory Index (AMI; Wiggins & Brandt, 1988), the Rey 15-Item Test (RIT; Rey, 1964), and the Mental Control Tests (MCT; Wechsler 1997b). The authors sought to address the following three hypotheses.

- (1) Simulators will perform significantly worse than non-simulators and head-injured patients on all four tasks.
- (2) A combination of tests will detect a higher proportion of simulators over non-simulators compared with one test alone.
- (3) All tests will show high sensitivity and specificity to detecting simulating malingering when compared with head-injured patients and non-simulating malingering participants.

## Method

### Participants

A total of 120 participants were taken from a volunteer community sample ( $N = 80$ ) and ( $N = 40$ ) participants with head injuries. Inclusion criteria were that volunteers were aged between 18 and 65 and had not previously had a head injury. Staff employed at the Walton Centre and the local office of the Health and Safety Executive were approached and asked to volunteer. Patients were recruited from those attending the local Brain Injury Rehabilitation Unit and the Transitional Rehabilitation Unit. Patients

also had to be aged between 18 and 65 years and had suffered a head injury, diagnosed independently of the study. All patients in the study had complained of memory problems and had completed neuropsychological assessments, which indicated significant memory impairments based on the Wechsler Memory Scale - Revised or a similar memory tests. The volunteer and head injury participants were matched for age and sex.

### **Measures and procedure**

#### *The Autobiographical Memory Index (AMI)*

The AMI is an 11-item test that asks for general information about the person being assessed. The questions are scored as plausible or incorrect. Incorrect answers were defined as clear examples of a wrong answer or an indication that the person had no knowledge of the answer. Wiggins and Brandt (1988) devised the AMI specifically to investigate malingering. People with acquired neurological disorders and head injuries are usually able to successfully complete 10 out of the 11 index questions. Cochrane *et al.* (1998) found that the AMI correctly differentiated between simulators and non-simulating participants.

#### *Rey 15-Item Memory Test (RIT)*

The Rey 15-item test has 15 items spread out across five separate lines containing three characters each. Each line consists of numbers, letters, or symbols. Participants were told that they would be shown a piece of A4 paper, which would have 15 different items on it. They had only 10 seconds to look at the A4 piece of paper with the 15 different items to remember. They were shown for 10 seconds and then asked to draw as many of the items as they could remember. Research has suggested a cut off of nine items or below would be appropriate for the detection of malingering (Goldberg & Millar 1986; Lezak, 1995). However, Bernard and Fowler (1990) indicated that a more appropriate cut off point is eight or below.

#### *Coin in the Hand Test (CIH)*

In this task participants were asked to remember in which hand a coin was placed. The experimenter showed a coin for approximately 2 seconds in one of his hands, and then put both his hands behind his back. The participants were then asked to count backward from 10 to 1 with their eyes closed. The experimenter then asked the subject to open their eyes and point to the hand holding the coin in. The experimenter then opened their hand and showed the subject the coin. The test was repeated for 10 trials and the order of hands was randomized equally between each hand. Kapur (1994) provided evidence of the validity of the test with five patients who had recognized neurological conditions and two suspected malingerers. The patients were correct on all trials whereas the suspected malingerers performed at best at chance level. Cochrane and colleagues (Cochrane *et al.*, 1998) found the CIH identified successfully 95% of simulators and Hanley *et al.* (1999) reported similar results with simulators obtaining a mean score of 4.10, whereas amnesics mean score was 9.6.

#### *Mental Control Test (MCT)*

The Mental Control Test is an optional test of the Wechsler Memory Scale - Third Edition, and incorporates eight tasks scored according to the accuracy and the subject's response. A combination of these scores gives the total score achieved for each subtest and a total score for all eight tests can be computed. The authors

decided to incorporate the mental control task as a result of a previous pilot that highlighted that subjects asked to feign malingering produced extremely poor results on this task, characterized by unusual mistakes and significantly impaired speed. This test had not previously been used to detect malingering although feigning could be determined in one of three ways: (1) the subject can reduce the speed of time taken to complete each individual task, (2) the subject can feign the accuracy of their responses, (3) the subject can utilize both 1 and 2.

### **Selection of volunteers to experimental groups**

At the start of the experiment, healthy volunteers were randomly allocated to the simulator or non-simulator group. Each group received one of two sets of instructions, which related to a simulation or a non-simulation scenario. The scenarios read as follows.

#### ***Simulator scenario***

'We want you to imagine that a year ago you were involved in a car crash in which you were a passenger. You received serious injuries to the head and were knocked unconscious for several hours, after which you were dazed and not taking things in properly for about 2 days. You were taken to hospital as an emergency and although you did not need surgery, the doctors were sufficiently concerned that you were admitted for a week during which you were in considerable pain. After you were discharged you continued to feel quite ill for several weeks, and you have not felt up to returning to work since then.

After the accident you decided to make a claim for the injuries suffered and you now have a solicitor who is dealing with the claim. People who have had a head injury commonly suffer from psychological problems such as forgetfulness and loss of concentration, and they have difficulty reasoning and thinking clearly. They may also suffer from emotional problems such as feelings of depression and anxiety as well as medical complaints, for example, affecting their eyesight or hearing. You know that the more affected you appear to be, the more financial compensation you are likely to receive. It is therefore in your best interests to magnify your complaints in order to obtain what you feel is rightly due as a result of being injured, but not so it is too obvious to the person testing you.

You have now been referred by your Solicitor to a psychologist who will attempt to evaluate the effect that the head injury has had upon you, and as part of the assessment you are asked to complete the following tests.'

#### ***Non-simulator scenario***

'Within the next 2-3 minutes the researcher will return to the room, and give you four short tests to carry out. We would like you to follow the instructions as closely as possible, and to complete the tests to the best of your ability. Please return this piece of paper to the envelope before the researcher returns. Do not mention the instructions given to you in this scenario until the researcher asks you after the four tests.'

### **Testing procedure for each group**

#### ***Head-injured patients (HI) (N = 40)***

Volunteers for the head-injured group were invited to assist in the study and were administered the tests using the standard format. All patients were asked to complete the four tests of memory to the best of their ability. The order of testing was AMI, RIT, CIH, and MCT.

**Non-simulators (NS; N = 40)**

Non-simulating volunteers were asked to complete the tests to the best of their ability. The order of the test was the same as for the head-injured patient group. No payment was made to the volunteers for their participation.

**Simulators (S; N = 40)**

Simulating volunteers were asked to complete the tests according to the malingering scenario (see above). The order of the test was the same as for the head-injured patient group. No payment was made to the volunteers for their participation.

**Data analysis**

The data were analysed using SPSS 11. Initial analysis was performed undertaking one-way between-participants ANOVA, followed by *post hoc* multiple comparison tests (Tukeys HSD) where appropriate. Receiver operator curve (ROC) analysis was applied to establish the sensitivity and specificity of each of the four tests to detect malingered performance.

**Results**

There were 120 participants of whom 69 were males with a mean age of 37 years. There were more males in the head injury group than the control groups (HI = 33, NS = 17, S = 19), but there were no differences between the groups for mean age (HI = 38.5, NS = 38, S = 37). Demographic details of the patients are presented in Table 1.

**Table 1.** Demographic details of the groups

	Simulators	Memory impaired	Non-simulating controls
Number	40	40	40
Sex: Male	19	33	17
Sex: Female	21	7	23
Mean age (years)	37	38.5	38
SD	12.7	13.5	11.5

Of the 40 patients in the memory-impaired group, 21 had suffered traumatic brain injuries, 3 had suffered encephalitis, 2 had suffered sub-arachnoid haemorrhages, and the remainder had experienced a range of organic pathologies. All of the patients had attended a head injury rehabilitation service and were considered to have significant memory problems. At the time of assessment, none were claiming for compensation.

When the groups performance on the four tests were analysed there were significant group effects. *Post hoc* multi-comparison tests revealed that the simulating group produced significantly lower scores ( $p < .01$ ) across all memory tests when compared with the non-simulating and head-injured group (see Table 2).

The utility of the individual tests to accurately classify group membership was examined. The authors employed a criterion cut-off point for establishing a member of the simulating group or non-simulating using the lower bond of the 95% confidence intervals for the head injured group across the four tests. Any participant who scored below this level was classified as a member of the simulator group. Results clearly

**Table 2.** Mean scores, standard deviations and *F* values obtained by each group for each test

Test	Head injury ( <i>N</i> = 40)	Non-simulator ( <i>N</i> = 40)	Simulator ( <i>N</i> = 40)	<i>F</i>	<i>p</i>
AMI Max = 11	10.68 (.65)	10.98 (.15)	9.05 (1.75)	36.47	<.01
RIT Max = 15	11.93 (2.83)	14.80 (.56)	8.43 (3.35)	62.52	<.01
CIH Max = 10	9.75 (.67)	10.00 (0)	6.53 (2.71)	142.37	<.01
MCT Max = 40	18 (4.95)	26.15 (3.88)	9.93 (5.98)	104.84	<.01

demonstrated that the CIH and the MCT tasks were the most successful in classifying group membership. Results are presented in Table 3.

**Table 3.** Mean scores for each group and the percentage of malingering simulators correctly classified for each of the four tests

Subject	AMI	RIT	CIH	MCT
Head injury	10.68	11.93	9.75	18
Non-simulators	10.98	14.80	10	26.15
Simulators	9.05	8.43	6.53	9.93
Misclassified as non-simulators	50%	32.5%	12.5%	17.5%
Correctly classified as a simulator	50%	67.5%	87.5%	82.5%

Using the CIH task alone allowed 87.5% of the Simulators to be correctly identified. Using a combination of the CIH and MCT correct classification of simulated performance was possible in all but three members of the simulated group, while misclassifying only two of the head-injured group as simulators (a 56 year-old man with multi-focal perfusion and IQ less than 80 and memory less than 70, and a 48 year old man with cerebral vascular disorder and memory greater than 70).

Performance by the simulated malingering participants showed higher sensitivity and specificity of two of the tests (CIH, MCT) than head injured and non-malingering participants. However, two tests (AMI, RIT) showed high sensitivity but low specificity when compared to the performance of the head injured and non-malingering participants (see Table 4).

**Table 4.** Sensitivity and specificity for differentiating between head-injured and malingering participants on the four tests

Comparison groups	Cut off score for detection of malingering <sup>a</sup>	Sensitivity	Specificity
AMI	> 10.50	77.5	72.5
RIT	> 8.00	77.5	65
CIH	> 8.50	92.5	87.5
MCT	> 13.50	80.5	82.5

<sup>a</sup>Any score of less than those presented would classify a participant as a simulator.

## Discussion

This paper set out to investigate the utility of four neuropsychological tests in detecting malingered performance. A high misclassification rate was found in two of the tests, AMI and RIT. An examination of the responses of the AMI revealed that there were only 6 of the 11 questions on which the simulators performed poorly. These included providing their mother's maiden name, what they had for breakfast, what they had for dinner, their phone number, age, and date of birth. Further research is underway that will investigate using a short form of this test.

The results of the RIT were equally poor in misclassifying both head-injured patients and simulators. A modification of the RIT (Paul, Franzen, Fremouw, & Cohen, 1992), removing the geometric shapes and adding an additional item to each row of three, led to increased specificity but variable sensitivity (Iverson & Franzen, 1996), leading Rogers (1997) to conclude that the utility of the RIT as a test of malingering on its own is limited. Cochrane *et al.* (1998), using a group of 20 simulators, found similar results. Modifications of the RIT reducing it to a 12-item measure have been made, although evidence of its sensitivity and specificity has yet to be documented.

In this study the RIT had a classification rate of only 67.5%. In a previous study by Millis and Kler (1995) the authors reported a classification rate of 57.5% amongst a group of suspected simulators. This finding is consistent with previous research (Arnett, Hammekke, & Schwartz, 1995). While on face value this test appears to be a reasonable measure for screening malingering, the high misclassification rate prevents it alone being a reliable and valid measure for detecting people who are likely to be malingering.

We chose to use the full version of the RIT to examine its performance with the groups in this particular design. In retrospect we might have gone directly to the shortened version, however, the authors were inclined to see how the older version performed in this design before applying the adapted version. Further research by the authors has utilized the shortened version of the RIT.

The CIH and MCT in combination appear to be very effective as screening tools for identifying obvious malingering. The CIH is a very simple task and is clearly very useful as a measure for screening obvious attempts to malingering, however for improving accuracy it is clearly important to include more than one task. The MCT has not been previously used as a screening tool, but the results from this study clearly merit its use as part of a battery for screening for malingering. Importantly, both tests are simple and brief to administer.

## Limitations to the study

The authors recognize that there are a number of limitations to the study that may affect the generalizability of the results, and these include the following.

- (1) The ethical difficulties of being unable to recruit patients who are pursuing claims for compensation.
- (2) The failure to consider randomizing the order of the test presentation to determine an order effect, which has been addressed in a further study.
- (3) The lack of the existence of a test that could be considered a 'gold standard' for detecting malingering that this battery of tests could be compared with.
- (4) The authors cannot be fully confident that the control participants would have

behaved similarly if they were actually pursuing a claim for compensation. It has been previously argued that simulators are unable to fully simulate as they do not have the incentive to motivate themselves into to the same situation as a true simulator would have with a litigation case impending. However, researchers who have incorporated a financial incentive to malingering have found no difference between incentive performers and simulators (Haines & Norris, 1995).

- (5) We applied our cut-off for group membership on the basis of the distribution of scores based on the lower 95% confidence interval for the head injured group. This approach maybe open to criticism but the authors feel that under the circumstances, it is a reasonable approach.
- (6) There are a number of issues surrounding the ecological validity of tests of malingering, particularly the choice of subjects and the use of a simulated design. These have been discussed previously (Rogers, 1997).

Despite these limitations, the findings are useful in providing evidence of the sensitivity and specificity of four measures commonly used and demonstrates the value of utilizing a combination of tests with high specificity and sensitivity rather than a single measure (Kapur, 1994). The CIH and MCT are easy to administer and take less than 5 minutes to complete. They will be useful to clinicians who are undertaking a legal assessment and who have doubts about the validity of symptoms being presented. The authors recommend that the tests themselves are only screening tools and should be supplemented with other clinical information before the clinicians decides on the balance of probability as to whether a patient is malingering or not.

## References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th edition)*. Washington, DC: American Psychiatric Association.
- Arnett, P. A., Hammekke, T. A., & Schwartz, L. (1995). Quantitative and qualitative performance on Rey's 15 Item Test in neurological patients and dissimulators. *Clinical Neuropsychologist*, *9*, 241-244.
- Baker, G. A., Hanley, J. R., Jackson, H. F., Kimmance, S., & Slade, P. (1993). Detecting the Faking of Amnesia: Performance differences between simulators and patients with memory impairment. *Journal of Clinical and Experimental Neuropsychology*, *15*, 668-684.
- Bernard, L. C., & Fowler, W. (1990). Assessing the validity of memory complaints: Performance of brain damaged and normal individuals on the Rey's Task to Detect Malingering. *Journal of Clinical Psychology*, *46*, 432-436.
- Binder, L. M., & Willis, S. C. (1991). Assessment of motivation after financially compensable minor head trauma. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*, 175-181.
- Brandt, J. (1988). Malingered amnesia. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 65-83). New York: Guilford Press.
- Cochrane, H. J., Baker, G. A., & Meudell, P. R. (1998). Simulating a memory impairment: Can amnesiacs implicitly outperform simulators. *British Journal of Clinical Psychology*, *37*, 31-48.
- Franzen, M. D., Iverson, G. L., & McCracken, L. M. (1990). The detection of malingering in neuropsychological assessments. *Neuropsychological Review*, *1*, 247-279.
- Goldberg, J. O., & Miller, H. R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology*, *42*, 792-795.



- Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesiacs do not forget. *Journal of Experimental Psychology: Human Learning and Memory*, *10*, 164-178.
- Haines, M. E., & Norris, M. P. (1995). Detecting the malingering of cognitive deficits: An update. *Neuropsychology review*, *5*(2), 125-148.
- Hanley, J. R., Baker, G. A., & Ledson, S. (1999). Detecting the faking of amnesia: A comparison of the effectiveness of three different techniques for distinguishing simulators from patients with amnesia. *Journal of Clinical and Experimental Neuropsychology*, *21*(1), 56-69.
- Hutchinson, L. (2001). *Disorders of simulation*. Madison, US: Psychosocial press.
- Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology*, *18*, 38-51.
- Kapur, N. (1994). The Coin in the Hand Test: A new bedside test for the detection of malingering in patients with suspected memory disorder. *Journal of Neurology, Neurosurgery and Psychiatry*, *57*, 385-386.
- Lezak, M. D. (1995). *Neuropsychological assessment*. Oxford: Oxford University Press.
- Millis, S. R., & Kler, S. (1995). Limitations of the Rey 15 Item Test in the detection of malingering. *Clinical Neuropsychologist*, *9*(3), 241-244.
- Pankratz, L. (1983). A new technique for the assessment and modification of feigned memory deficits. *Perceptual and Motor skills*, *57*, 367-372.
- Paul, D., Franzen, M. D., Fremouw, W., & Cohen, S. (1992). Standardisation and validation of two tests used to detect malingering. *International Journal of Clinical Neuropsychology*, *14*, 1-9.
- Resnik, P. J. (1984). The detection of malingered mental illness. *Behavioural Sciences and the Law*, *2*, 21-38.
- Rey, A. (1964). *L'examen clinique en Psychologie (The Clinical Examination in Psychology)*. Paris: Presses Universitaires de France.
- Rogers, R. (1997). *Clinical assessment of malingering and deception*. New York: Guildford Press.
- Rogers, R., Harrell, E. H., & Liff, C. D. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review*, *13*, 255-274.
- Schacter, D. L. (1986). On the relation between genuine and simulated amnesia. *Behavioural Sciences and the Law*, *4*, 47-64.
- Wechsler, D. (1997b). WMS - III Wechsler Memory Scale (3rd ed.) San Antonio; The Psychological Corporation.
- Wiggins, E., & Brandt, J. (1988). The detection of simulated amnesia. *Law and Human Behaviour*, *12*, 57-78.

Received 18 July 2003; revised version received 12 March 2004