# Malingering on the RAVLT part II: Detection strategies.

Karen Sullivan, PhD, Cassandra Deffenti, & Beth Keane
Queensland University of Technology
AUSTRALIA

Correspondence concerning this article should be addressed to Karen Sullivan at the School of Psychology and Counselling, Queensland University of Technology, Carseldine, Queensland 4034, Australia. Electronic mail should be sent to ka.sullivan@qut.edu.au

Keywords: neuropsychology, malingering, RAVLT, serial position effect

Abstract

In this study two potential indices of malingering derived from the Rey Auditory Verbal Learning Test (RAVLT) were evaluated as a means of detecting malingering.  These were indices based on: discrepancies between recognition-recall scores, and differences in the serial position effect (SPE).  Sixty undergraduate students were randomly assigned to one of four conditions: malingerers, malingerers-with-warnings, warning-only, and control.  Incentives were offered to participants in all conditions to encourage faking in a believable manner (malingering conditions), or to encourage optimal performance (non-malingering conditions). Two predictions were made.  First, it was predicted that the serial position curve for subjects in malingering conditions would show suppression of primacy effects relative to non-malingerers.  Second, it was predicted that recall would be better than recognition for subjects in malingering but not non-malingerers conditions.  The utility of these indices was also explored in the context of providing subjects' with warnings regarding use of methods to detect malingering.  Results indicated that both indices failed to reliably differentiate between malingerers and non-malingerers, and warnings failed to modify participants' behaviour.

Malingering on the RAVLT part II: Detection strategies.

There is a growing body of literature reflecting attempts to identify and improve malingering detection strategies in a neuropsychological context (see Heubrock & Petermann, 1998 for a recent review). For example, methods of detection that have been investigated previously include attempts to identify discrepancies in information provided at interview, during testing, and through observation, as well as attempts to examine test scores to determine the presence of atypical patterns of performance (Haines & Norris, 1995; Rogers, Harrell, & Liff, 1993). Atypical patterns of performance that have been investigated for their potential to distinguish malingerers from non-malingerers include: below-chance performance, and performance that deviates from patterns produced by reliable cognitive phenomenon (e.g., absence of priming-effects, serial-position-effects, or higher rates of recall than recognition on memory tasks). This paper focuses on the absence of serial position effects and atypical discrepancies between recognition-recall as a means of detecting malingering.

Serial Position Effects

Serial position effects (SPE) are typically found on verbal list learning tasks and describe a pattern of results that can be consistently demonstrated in healthy adults (Haberlandt, 1997). That is, when asked to recall a list of words, the first- and last-third of a list are more likely to be remembered than words from the middle of the list. The increased likelihood of recall of words from these positions produces a U-shaped pattern of results that is indicative of primacy and recency effects.

Differences in SPEs have been explored as a means of detecting malingering previously. For example, Bernard (1991) showed that SPEs differentiated malingerers from non-malingerers on the Rey Auditory Verbal Learning (RAVLT). Malingerers were found to suppress recall of words in the first-third of the list, demonstrating significantly reduced primacy effects relative to non-malingerers. The rationale for exploring the SPE as a tool for detecting malingering is based on the nature of this phenomenon. That is, as an index of implicit memory function, serial position curves usually occur beyond the awareness of the person completing the memory task and can be reliably demonstrated. Therefore, the absence of SPEs might indicate an attempt to consciously modify responses, producing an atypical pattern of results.

Studies attempting to replicate Bernard's (1991) finding of reduced primacy effects among malingerers have yielded contradictory results. For example, a later study conducted by Bernard and co-workers found no significant differences in SPEs between simulators and controls (Bernard, Houston, & Natoli, 1993). Although malingerers performed significantly worse than controls on the RAVLT, results from both groups conformed to the U-shaped curve that is characteristic of normal primacy and recency effects. Similarly, the results of Flowers, Sheridan, and Shadbolt (1996) showed that serial position curves for malingerers and non-malingerers had the same shape, and groups could not be differentiated based on a comparison of SPEs. Given this inconsistent pattern of results, it is unclear to what the analysis of SPEs can be used to detect malingering, or more specifically, to identify deliberately poor performance.

There may also be some contexts in which analysis of SPEs is more effective than others, although this has not been investigated previously. For example, it may be possible to demonstrate the utility of an SPE-based index in contexts where a warning about malingering is explicitly provided. The warning literature suggests that would-be malingerers, when informed that methods of detecting malingering will be employed during assessment, modify their behaviour and do not perform as poorly as "unwarned" malingerers (Johnson & Lesniak-Karpiak, 1997). A possible explanation for this result lies in deterrence theory (Sullivan, Keane, & Deffenti, in press). However, it has also been suggested that warnings may result in more sophisticated attempts to malinger (Youngjohn, Lees-Haley, & Binder, 1999). This suggestion is based around concerns that rather than deterring would-be malingerers from faking or exaggerating deficits, warnings may produce more sophisticated malingerers who are more difficult to detect. If this is the case and the provision of warnings produces more "sophisticated" malingerers, analysis of differences in SPEs (or other indices of atypical performance that are usually beyond the subject's awareness) might enable there to be detected.

Recognition-recall

Another approach to the detection of malingerers based on atypical patterns of performance involves use recognition-recall discrepancy scores. For example, it is generally accepted that healthy people recall less than they recognise (Bernard, 1990) and importantly, this pattern of results also usually occurs beyond the subject's awareness. Further, the results of several malingering studies suggest that malingering and non-malingering subjects may be distinguished by their performance on recognition versus recall tasks (Bernard et al., 1993; Flowers et al., 1996; Greiffenstien, Baker, & Gola, 1994). Specifically, Bernard and colleagues found that malingerers performed at a level that was disproportionately poor on recognition tasks, relative to free-recall tests. Similarly, Flowers and colleagues (1996) showed that malingerers identified fewer words than they recalled.

The reason why malingerers perform at a level that is disproportionately poor on recognition tasks, is thought to be due to the relative ease of recognition memory tasks, since this limits options for malingering (Bernard et al., 1993; Johnson & Lesniak-Karpiak, 1997). For example, on recognition memory tasks malingerers have one basic strategy: omit previously recalled words. The notion that tasks with limited faking options may provide a useful means of detecting malingering also underlies the effectiveness of two-alternative forced-choice symptom validity tests (Binder & Pankratz, 1987; Pankratz, 1983; Pankratz, Fausti, & Peed, 1975). Overall, while there appears to be some consensus in the literature that performance of malingerers on recognition tasks may be disproportionately affected compared to free recall tasks, the effect of warnings on this pattern of results has yet to be established.

In summary, there is clearly a need for further investigation of the utility of analyses based on atypical patterns of performance, such as the absence of SPEs and atypical recognition-recall discrepancies as a means of detecting malingering. Therefore, the aims of this study were two-fold. First, to expand our understanding of the effects of SPE as a means of detecting malingering. Second, to further investigate the utility of recognition-recall discrepancy scores as method of detecting malingerers. In addition, the effect of warnings on these indices was also explored.

To achieve these aims, two hypotheses were proposed. First, based on the results of Bernard (1991), it was expected that serial position curves would differentiate malingerers from non-malingerers (warning-only and control), with malingerers (malingerers-with-warning, and malingerers) producing a reduced primacy effect. This result would support the proposition that analysis of SPEs may be a useful additional tool for detecting malingerers on list learning tasks. It would also support the proposition that warnings do not alter the behaviour of malingerers, at least on SPE indices, which would be contrary to expectations of Youngjohn and colleagues (Youngjohn et al., 1999).

Second, it was expected that malingerers (malingerers and malingerers-with-warning) would perform worse on the recognition trial of the RAVLT than free recall trials, and that this would distinguish malingerers from non-malingerers. This result would suggest that poorer performance on recognition tasks may raise suspicion of malingering. It would also suggest that warnings do not alter malingering strategies.

Method

Participants

Participants were recruited from the School of Psychology and Counselling undergraduate subject pool and the Law Faculty, at the Queensland University of Technology. Sixty-one volunteers participated in this study. The demographics of the sample and exclusion criteria for this study have been described fully elsewhere (Sullivan et al., in press). Briefly, age ranged from 17 to 54, and a one-way ANOVA found no significant differences between groups on age, $F(3, 56) = 1.01$, $p > .05$.

Subjects were excluded from this study if they had experienced brain injury, concussion, and amnesia or memory loss, or if they were unable to understand the instructions provided, as measured by the Intervention Effectiveness Appraisal Questionnaire (IEA). Using these criteria, data from one subject was excluded from analysis on the basis of past medical history. No responses were excluded from analysis based on responses to the IEA which was used in this study as a measure of subject compliance. That is, no responses were excluded due to failure to understand the instructions for this experiment, including instructions on how to malinger.

Design

Hypotheses regarding the detection of malingering using analysis of atypical patterns of performance were examined in this study using mixed ANOVAs. That is, serial position differences between groups were examined using a three-way mixed ANOVA with two between-groups variables (group and occasion), and one within-groups variable (serial position with three levels). The dependent variable for this analysis was the total number of correctly recalled words in each third of the RAVLT. That is, serial position data was obtained by adding the occurrence of correctly identified words for the first- (words 1-5), middle- (words 6-10), and the last-third (words 11-15) of the RAVLT across five trials. In

subsequent sections of this paper these scores are referred to as primacy, mid-list, and recency scores respectively.

Differences between recognition and recall scores were examined using a two-way mixed ANOVA with one between-groups variable (group) and one within-groups variable (occasion). The dependent variable for this analysis was the recognition-recall discrepancy score, which was calculated following the method used by Flowers and colleagues (1996). That is, recognition scores were subtracted from the highest recall score obtained over trials one to five.

Materials

The RAVLT (or AVLT as it is also known; Spreen & Strauss, 1998) is a 15-item list learning task with five learning trials. Between each trial, the target list is re-read to subjects. There is an interference list (a new list of 15 words) that is read to the subject after the fifth learning trial, and after recall of the interference list is attempted, recall of the original list is undertaken. The RAVLT also has a recognition trial in which subjects are asked to identify words from the first 15-item list using a standard target-distracter word set. This test was administered and scored in accordance with standardised instructions (see Spreen & Strauss, 1998).

Indices derived from the RAVLT were used as dependent measures in this study for several reasons. First, the RAVLT is widely recognised as a valid measure of memory (Rosenberg, Ryan, & Prifitera, 1984; Spreen & Strauss, 1998). Second, the RAVLT has been used previously in published studies of malingering (King, Gfeller, & Davis, 1998). Third, the RAVLT is a popular test that is frequently used in clinical settings (Spreen & Strauss, 1998; Sullivan & Bowden, 1997). Fourth, the RAVLT has alternate forms with acceptable alternate form reliability, which is important when using repeated-measures designs (Spreen & Strauss, 1998). Fifth, and perhaps most importantly, as Flowers and colleagues (1996) have illustrated, there are a range of scores that can be derived from the RAVLT that may prove useful as indicators of atypical patterns of performance.

Procedure

The procedure for this study has been described in detail elsewhere (Sullivan et al., in press). Briefly, this study included three stages: pre-intervention, intervention, and post-intervention. Two examiners were involved in the testing of each subject to create a blind examiner: one examiner facilitated the pre-intervention and intervention stages, the other examiner administered the RAVLT during the post-intervention stage. At the pre-intervention stage, participants signed a letter of consent, and then completed the RAVLT and a demographics questionnaire.

At the intervention stage, participants were randomly allocated to groups by handing them a sealed envelope containing instructions specific to one of the four conditions (malingerers, malingerers-with-warning, warning-only, and control; 15 subjects per group). The only difference between the groups was the instructions they received regarding how to perform on the RAVLT during the post-intervention stage. Participants were encouraged to read the

instructions for their condition, and ask any questions they may have with pre-intervention examiner only (ie no questions were to be asked regarding intervention instructions of the examiner administering the follow-up RAVLT). All four groups watched a video of memory deficits experienced by some people with acquired brain injury. All groups were told they would receive a reward for performing to the best of their ability, or in a credible (undetectable) manner, depending on group allocation. Rewards in this study were offered in each of the four conditions, and although rewards were presented in such a way as to appear differential to subjects (i.e., rewards appeared to depend on how well participants performed, or whether they faked deficits credibly), each individual received the reward regardless of performance. The nature of the reward was not disclosed to subjects until completion of the task and consisted of five dollars of photocopy credit.

At post-intervention, participants were required to complete an alternate form of the RAVLT in the manner determined by the instructions they had received. For example, participants in the malingering group were asked to complete the RAVLT in a way that reflected difficulties they had observed on the video but appeared believable and realistic. In the malingering-with-warning group, the same instructions were issued but additional information was provided as follows: "Warning: efforts to malinger (fake or exaggerate difficulties) on the memory test will be detected through in-built methods within this test".

Results

To determine the overall level of performance of subjects in this study and provide a gross measure of equivalence across groups, descriptive statistics for RAVLT occasion one trial scores (pre-intervention) were calculated. These statistics showed a consistent increase in free recall scores across trials for all groups, with performance in the average range compared to age-corrected norms (Spreen & Strauss, 1998).

Analysis of Serial Position Effects.

A three-way mixed repeated measures ANOVA was conducted to determine whether primacy, recency, or mid-list scores differed between groups on occasions one and two. The results of this analysis revealed significant main effects for serial position $\underline{F}$ (2,56) = 43.02, $\underline{p}$ = .000, group $\underline{F}$ (3,56) = 12.11, $\underline{p}$ = .000, and occasion $\underline{F}$ (1,56) = 98.71, $\underline{p}$ = .000. Significant interactions were found between occasion and group, $\underline{F}$ (3,56) = 27.19, $\underline{p}$ =.000, and occasion and serial position, $\underline{F}$ (2,56) = 14.78, $\underline{p}$ = .000, but not for serial position and group, $\underline{F}$ (6,56) = 1.63, $\underline{p}$ = .146, or the three-way interaction between occasion, serial position and group, $\underline{F}$ (3,56) = 1.49, $\underline{p}$ = .189.

Significant interaction effects involving serial position were subject to further analysis. That is, the significant occasion by position interaction effect was investigated by calculating the difference between serial position on testing occasions one and two and conducting a one-way ANOVA on difference scores. Significant differences were found in all positions: primacy $\underline{F}$ (3, 56) = 15.08, $\underline{p}$= .000, mid-list $\underline{F}$ (3, 56) = 17.10, $\underline{p}$= .000, and recency $\underline{F}$ (3, 56) = 7.73, $\underline{p}$= .000. With alpha set at .05, Tukey HSD post-hoc comparisons revealed significant differences between controls and both malingering groups in each of the three positions. No significant differences were found between non-malingering groups (ie. control and warning-only), or

malingering groups (ie. malingerers and malingerers-with-warning. These results are illustrated in Figure 1. For clarity, pre- and post-intervention results are shown in panels (a) and (b) of Figure 1 respectively. Figure 1a shows a reasonably consistent pattern of results for all groups, in which primacy and recency effects are apparent. Figure 1b shows that the shape of the curves obtained post-intervention for each group was similar, although shifted downwards in malingering conditions. That is, primacy and recency effects remain apparent, but there is a general reduction in the number of words recalled across all thirds of the list in malingering conditions.

1 (a)
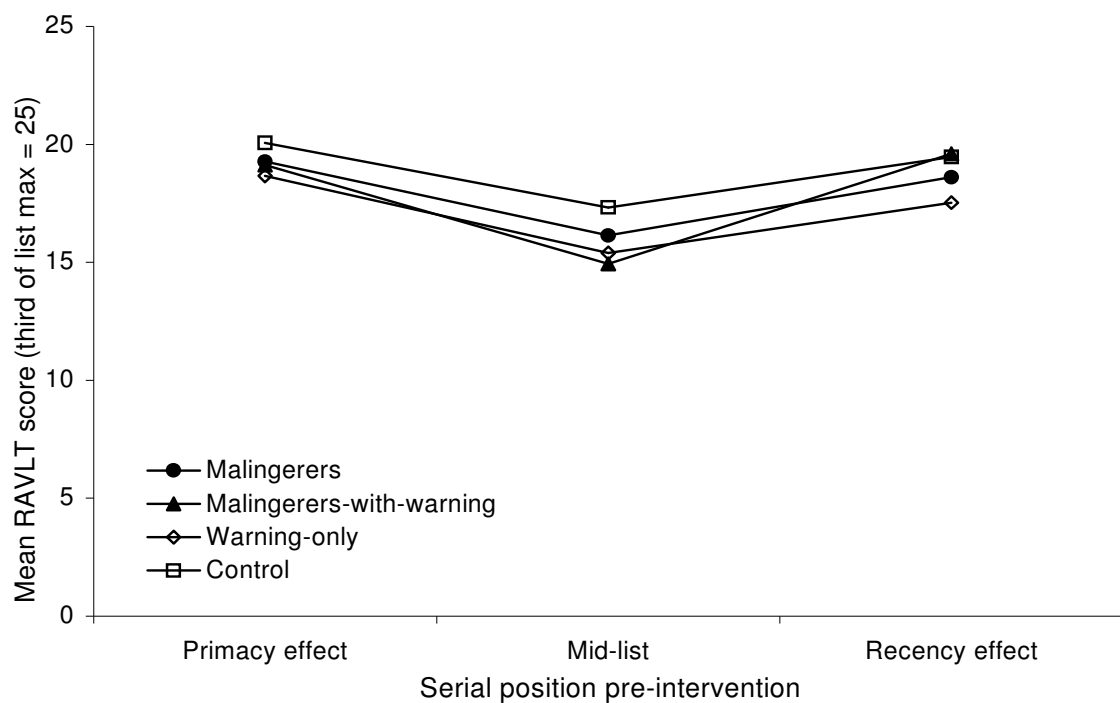


1 (b)
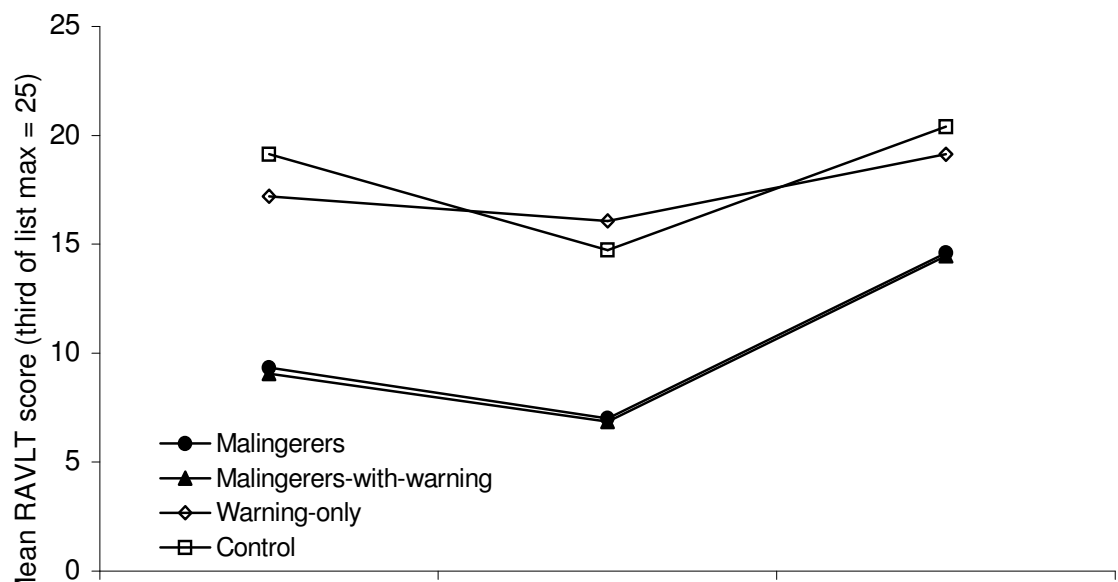
<u>Figure 1</u> Primacy, mid-list and recency effects for all groups pre- (a) and post-intervention (b). Mean RAVLT scores represent scores across list thirds (e.g., trials 1 to 5 = primacy effect)

<u>Recognition and Recall Comparisons</u>

A two-way mixed repeated measures ANOVA exploring differences in discrepancy scores revealed a significant effect for occasion $F (1,56) = 10.57$, $p = .002$, but not group $F (3,56) = .61$, $p = .610$, or the interaction between discrepancy scores and group $F (3,56) = 2.67$, $p = .056$. Figure 2 shows the differences between recognition and highest recall scores for all groups. For clarity, data is presented separately for occasion one (pre-intervention) and occasion two (post-intervention) in parts (a) and (b) of Figure 2 respectively. Figure 2a shows that a relatively consistent pattern of results across groups pre-intervention. That is, recognition scores (filled bars) are close to ceiling and, on average, less than one word greater than the highest immediate recall score (unfilled bars). Figure 2b shows a change in patterns for malingering groups, but not for warning-only and control groups where scores on recognition and recall tasks remain close to ceiling. For example, post-intervention, there is a sharp decline in both recognition and recall scores of malingerers and malingerers-with-warning, relative non-malingering conditions (see Figure 2(b)). The pattern of differences found suggests a general reduction in the highest number of words recalled and the number of words correctly recognized for malingering groups compared to non-malingerers.

2(a)



Recognition and highest recall scores pre-intervention
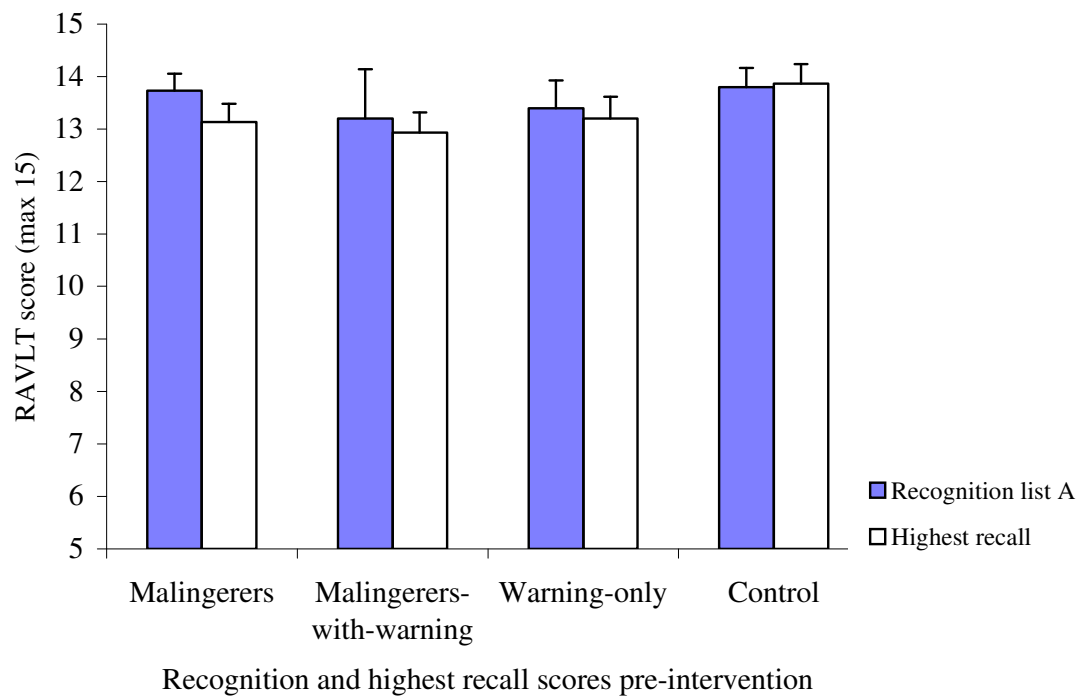
2(b)
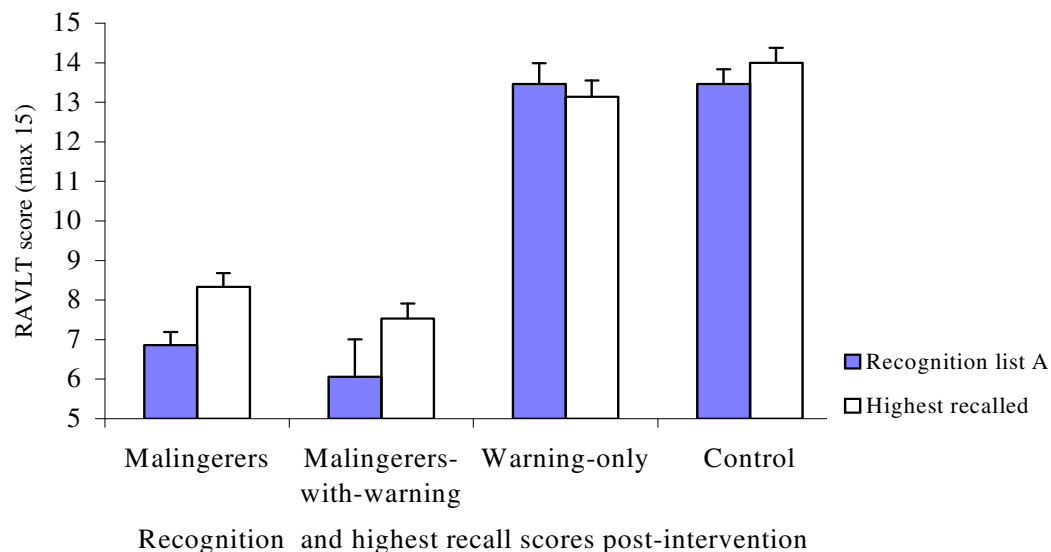


Figure 2  Mean list A recognition score and highest immediate recall trial score (list A) across all groups on testing occasion one (2a) and two (2b).  Standard error bars for all groups are shown.

Discussion

The aim of the current study was to further investigate the effectiveness of atypical patterns of performance on the RAVLT as a means of detecting malingering.  Specifically, this study aimed to determine whether analysis of differences in SPE or discrepancies in recognition-recall scores might be a useful means of detecting malingering.  The effectiveness of these indices in a setting where warnings about malingering were explicitly provided was also explored.

To achieve these aims, two hypotheses were proposed.  First, the hypothesis that serial position curves would differentiate between groups, with malingerers producing a reduced primacy effect, was not supported.  That is, it was expected that the serial position curve would differentiate malingerers from non-malingerers.  In addition it was proposed that malingerers would have a reduced primacy effect (fewer words remembered from the first third of the list).  However, analysis of the serial position curve revealed that malingerers performed at lower levels than non-malingerers across all thirds of the list, consequently the U-shaped curve was maintained.  In particular, there was no significant reduction in primacy scores for any group post-intervention.  This result suggests that reduced primacy effects may not be a reliable indicator of malingering, or more specifically, may not help identify deliberate attempts to perform poorly on the RAVLT.

Previous research on the utility of the serial position curve as a means of detecting malingering has produced inconsistent results, as noted previously. Contrary to expectations, we failed to find evidence of a reduced primacy effect in malingerers, as had been reported earlier by Bernard (1991). This inconsistency is unlikely to be due to differences in the properties of measures used since both studies used the RAVLT. However, there are several other studies that have also failed to show that suppression of primacy effects can differentiate between malingerers and non-malingerers (Bernard et al., 1993; Flowers et al., 1996). Overall, the results of studies failing to show the suppression effect suggest that whilst the shape of the serial position curve is maintained, there is a downward shift for malingerers indicating a general decrease in the number of words recalled, regardless of word position.

The second hypothesis proposed for this study was that malingerers would perform poorer on the recognition task relative to recall and that this would differentiate the groups, however this hypothesis was also not supported. That is, results indicated that malingerers and non-malingerers could not be differentiated on the basis of the recognition-recall discrepancy scores. The results of recognition-recall comparisons in this study were not consistent with results from previous research which suggests that malingerers perform at a level that is disproportionately poor on forced-choice recognition memory tasks (e.g., Bernard, 1990; Bernard et al., 1993; Flowers et al., 1996; Greiffenstien et al., 1994; Wiggins & Brandt, 1988). In this study, results suggest that performance on both recognition and recall tasks was reduced in malingering conditions and there was no significant differential effect depending on task type. A trend in the expected direction was apparent however (ie. recall greater than recognition in both malingering conditions; see Figure 2), which suggests that the magnitude of the reduction in recall on both types of tasks may have masked any differential effect in this study.

The failure to discriminate between malingerers and malingerers-with-warning using analyses based on atypical RAVLT performance in this study, is interesting in light of recent suggestions that warning malingerers about the possibility of detection may produce more sophisticated attempts to malinger (Youngjohn et al., 1999). For example, if warning malingerers about the possibility of detection encourages malingerers to refine their strategy for faking, it may be that these indices could be used to detect performance differences when warnings are provided. However, the results of this study suggest that neither analysis of the SPE, or analysis of recognition-recall discrepancy scores, revealed a different pattern of results among warned and unwarned malingerers. Further, analyses based on total RAVLT scores also failed to discriminate between these groups (see Sullivan et al., in press). It is possible there may be other ways of analysing this data that could reveal subtle differences in the strategies used by malingerers-with-warning, that are not apparent from the analyses conducted for this study.

In conclusion, we failed to find evidence of the effectiveness of recognition-recall scores, or SPE as a means of detecting malingering in simulators. It is important to note that, as with all simulation studies, there are issues relating the generalisability of these results to genuine malingerers (Haines & Norris, 1995; Nies & Sweet, 1994). As such, further studies may be needed to explore the utility of these indices (particularly recall-recognition score discrepancies) in a sample of known malingerers. This is important to determine whether

malingerers behave in the same way as simulators in this study before these indices are ruled out as ineffective.

References

American Psychiatric Association. (1994).  <u>Diagnostic and Statistical Manual of Mental Disorders,</u>(4<sup>th</sup> Ed.).  Washington: American Psychiatric Association.

Bernard, L. C. (1990).  Prospects for faking believable memory deficits on neuropsychological test and the use of incentives in simulation research.  <u>Journal of Clinical and Experimental Neuropsychology, 12,</u> 715-728.

Bernard, L. C. (1991).  The detection of faked deficits on the Rey Auditory Verbal Learning Test: The effects of serial position.  <u>Archives of Clinical Neuropsychology, 6,</u> 81-88

Bernard, L. C., Houston, W., & Natoli, L. (1993).  Malingering on neuropsychological memory tests: Potential objective indicators.  <u>Journal of Clinical Psychology, 49,</u> 45-53.

Binder, L. M., & Pankratz, L. (1987).  Neuropsychological evidence of a factious memory complaint.  <u>Journal of Clinical and Experimental Neuropsychology, 9,</u> 167-171.

Faust, D., & Guilmette, T. J. (1990).  To say it isn't so doesn't prove that it isn't: Research on the detection of malingering: Reply to Bigler.  <u>Journal of Consulting and Clinical Psychology, 58,</u> 248-250.

Faust, D., Hart, K., & Guilmette, T. J. (1988).  Pediatric malingering: The capacity of children to fake believable deficits on neuropsychological tests.  <u>Journal of Consulting and Clinical Psychology, 56,</u> 578-582.

Flowers, K. A., Sheridan, M. R., & Shadbolt, H. (1996).  Simulation of amnesia by normals on the Rey Auditory Verbal Learning Test.  <u>Journal of Neurolinguistics, 9,</u> 147-156.

Greiffenstien, M. F., Baker, W. J., & Gola, T. (1994).  Validation of malingered amnesia measures with large clinical samples.  <u>Journal of Psychological Assessment, 6,</u> 218-224.

Haberlandt, K. (1997).  <u>Cognitive psychology</u>  (2nd ed.).  Needham Heights, MA: Allyn & Bacon

Hayward, L., Hall, M., & Zurbrick, S. R. (1987).  Can localised brain impairment be simulated on neuropsychological test profiles.  <u>Australian and New Zealand Journal of Psychiatry, 21,</u> 87-93.

Heaton, R. K., Smith, H. H., Lehman, R. A. W., & Vogt, A. T. (1978).  Prospects for faking believable deficits on neuropsychological tests.  <u>Journal of Consulting and Clinical Psychology, 46,</u> 892-900.

Heubrock, D., & Petermann, F. (1998).  Neuropsychological assessment of suspected malingering: Research results, evaluation techniques, and further directions of research and application.  <u>European Journal of Psychological Assessment, 14,</u> 211-225.

Johnson, J. L., & Lesniak-Karpiak, K. (1997).  The effects of warning on malingering on memory and motor tasks in college samples.  <u>Archives of Clinical Neuropsychology, 12,</u> 231-238.

King, J. H., Gfeller, J. D., & Davis, H. P. (1998).  Detecting simulated memory impairment with the Rey Auditory Verbal Learning Test: Implications of base rates and study generalizability.  <u>Journal of Clinical and Experimental Neuropsychology, 20,</u> 603-612.

Nies, K. J., & Sweet, J. J. (1994).  Neuropsychological assessment and malingering: A critical review of past and present strategies.  <u>Archives of Clinical Neuropsychology, 9,</u> 501-552.

Pankratz, L. (1983).  A new technique for the assessment and modification of feigned memory deficits.  <u>Perceptual and Motor Skills, 57,</u> 367-372.

Pankratz, L., Fausti, S. A., & Peed, S. (1975).  Case study: A forced-choice technique to evaluate deafness in the hysterical and malingering patients.  Journal of Consulting and Counselling Psychology, 43, 421-422.

Rogers, R., & Cruise, K. R. (1998).  Assessment of malingering with simulation designs: Threats to external validity.  Law and Human Behaviour, 22, 273-285.

Rogers, R., Harrell, E., & Liff, C. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations.  Clinical Psychology Review,13, 255-274.

Rosenberg, S. J., Ryan, J. J., & Prifitera, A. (1984).  Rey Auditory Verbal Learning Test performance of patients with and without memory impairment.  Journal of Clinical Psychology, 40, 785-787.

Spreen, O., & Strauss, E. (1998).  A compendium of neuropsychological tests: Administration, norms and commentary.  New York: Oxford University Press.

Sullivan, K. A., & Bowden, S. C. (1997).  Which tests do neuropsychologists use? Journal of Clinical Psychology, 53, 657-661.

Sullivan, K. A., Keane, B., & Deffenti, C. (in press).  Malingering on the RAVLT part I: Deterrence strategies.  Archives of Clinical Neuropsychology.

Youngjohn, J. R., Lees-Haley, P. R., Binder, L. M. (1999).  Comment: Warning malingerers produces more sophisticated malingering.  Archives of Clinical Neuropsychology, 14 (6), 511-515.