The Effect of Item Type on Performance of the Matrix Reasoning Subtest of the WAIS-III in Traumatically Brain Injured and Non Brain-injured Control Participants

Simon F. Crowe and Robyn M. Bittner
La Trobe University, Australia

Ramona Raggl
La Trobe University and University of Southern Queensland, Australia

Graeme Senior
University of Southern Queensland, Australia

Qualitative analysis of neuropsychological instruments has been a long tradition in neuropsychological assessment. This study extended this type of analysis to the Matrix Reasoning (MR) subtest of the WAIS-III. The study compared the performance of TBI participants on the item types identified within the MR subtest (i.e., pattern completion, classification, analogy and serial reasoning) with a group of normal controls. MR items were classified into categories (as defined respectively by the Psychological Corporation and by our own research definition). Ninety-three non brain-injured control and 72 brain injured control participants were included in the study. One way analysis of variance indicated that the TBI group performed significantly worse than the non brain-injured group the MR performance overall as well as for both the Psychological Corporation classification and on the research defined categories. Within group analysis revealed that both groups performed significantly differently across the item categories with the most difficult categories being analogy and serial reasoning for the research defined categories and the classification and serial reasoning categories for the Psychological Corporation-defined groups. The results of the study indicate that an item type analysis of the MR performance may further contribute to the qualitative aspects of diagnostic formulation.

The Raven's Progressive Matrices (RPM) was originally designed by Raven (1948) to measure Spearman's *g* factor, and it is considered by many to be the "gold standard" of inductive reasoning tasks. Raven (1948) viewed the matrices test as a measure of a person's capacity to form comparisons, reason by analogy, and develop a logical method of thinking regardless of previously acquired information that can be used as a means of estimating innate "inductive ability".

The RPM has been employed in numerous neuropsychological investigations including developmental studies (i.e., with both young [Green & Kluever, 1991; Gutman, 1974] and elderly participants [Babcock, 1994; Salthouse, 1993]), and in studies of right versus left hemisphere

Address for correspondence: Associate Professor Simon F. Crowe, School of Psychological Science, La Trobe University, Bundoora, Australia, VIC 3086, Australia. Email: s.crowe@latrobe.edu.au

lesions (Soukup, Harrell, & Clark, 1994; Villardita, 1985).

Villardita (1985) has noted that brain impaired individuals perform differently on different aspects of the RPM. Set 1 of the matrices requires "identification of sameness" and posed particular problems for right brain lesioned patients. Set 2 involved the principle of symmetry and was selectively failed by aphasic patients; and Set 3 was more demanding on analogical reasoning and conceptual thinking and was performed poorly by all left brain damaged patients, aphasics as well as non-aphasics.

Early studies contended that the effect of impairment of RPM in right hemisphere strokes was due to visual neglect of the display. However, Soukup et al. (1994) noted no evidence for this when an alteration of the display was presented as a column of options in the right hemispace, indicating that the problem is cognitive rather than sensory in origin.

Numerous formulations have been proposed to account for what the RPM measures, including verbal versus nonverbal abilities; gestalt versus analytical skills; memory; spatial abilities; processing speed; and working memory (Carpenter, Just, & Shell, 1990).

In the non-metric scaling technique described by Marshalek, Lohman, and Snow (1983) the RPM was found to be at the centre of a group of measures thought to measure analytical intelligence. Tests in the centre of the group of measures (i.e., conducting necessary arithmetic operations, solving number, geometric and verbal analogies, determining letter series, paper folding and surface development) were thought to involve abstract reasoning, induction of relations and deduction. The source of the RPM's centrality according to Marshalek et al. (1983) is that more "complex tasks require more involvement of executive assembly and control processes that structure and analyze the problem, assemble a strategy of attack on it, monitor the performance process, and adapt these strategies as performance proceeds" (p. 124).

Salthouse (1993) conducted four studies to investigate how working memory (WM) mediates the age differences in RPM. He noted the statistical control of WM reduces the age-related variance in RPM performance by 70% indicating that age-related change in working memory functions are crucial to performance on the task. He noted that the factors responsible for the age differences were not the same as those contributing to item difficulty.

Raven (1948) has suggested that while the RPM assesses a unitary trait, five principles underlie the solution of the standard matrices: pattern completion, analogy, systematic pattern completion, systematic permutation and systematic resolution of figures into parts.

Recently a new subtest has been added to the Wechsler Adult Intelligence Scale-III (WAIS-III: Wechsler, 1997), in the form of the Matrix Reasoning subtest (MR). This was done with a view to increasing the WAIS's utility in evaluating novel, abstract and nonverbal reasoning abilities (Dugbartey, Sanchez, Rosenbaum, Mahurin, Davis, & Townes, 1999).

The MR subtest contains 26 items and is similar to other forms of matrix analogy task such as the RPM. The items require the examinee to use reasoning and problem-solving skills as well as performing mental manipulation with each item. The reliability coefficients of the task in the standardisation sample of the WAIS-III are quite stable across age groups (i.e., average r = .9) and the subtest correlates with Block Design and loads on the perceptual organisation index of the WAIS-III on factor analysis (Tulsky, Saklofske, & Zhu, 2003).

A recent comparison between patients with mild or moderately severe injury and matched control subjects undertaken by Donders, Tulsky and Zhu (2001) has indicated that the subtest is not particularly sensitive to the effects of traumatic brain injury (TBI). These investigators noted that "Matrix Reasoning, on the other hand, is not sensitive to the sequelae of TBI and more studies are needed to determine how it can be used for neuropsychological assessment purposes" (Donders et al., 2001, p. 892).

It is well known that TBI is often associated with deficits in speed of information processing (e.g., Dikmen, Machamer, Winn, & Temkin, 1995), hence an untimed task such as the MR subtest would be less likely to be affected by the effects of processing speed than would other performance subtests such as Block Design, Picture Completion or Object Assembly which give bonuses for speeded response.

The authors of the WAIS-III (Wechsler, 1997) have contended that there are four problem types in the MRT: pattern completion, classification, analogy and serial reasoning. To date, no examination of the relative sensitivity of these item types to cognitive impairment has been undertaken. It may be the case that a more fine-grained analysis of item type within the MR may reveal differences between TBI and non-injured controls which were not revealed by scaled score comparisons (Donders et al., 2001). The aim of this study was to investigate the performance of individuals who had suffered traumatic brain injury on the various item types of the MR and to compare these with a group of non-injured participants, to determine if these individuals showed differences in the pattern of performance on the different item types.

METHOD

Participants

The sample of 72 TBI participants (59 male, 13 female) ranged in age from 17 to 58 years (mean = 37.33, *SD* = 9.62) and in education from 7 to 18 years (mean = 11.23, *SD* = 2.16). All TBI participants were part of an archival database derived from a private clinical practice, and were a consecutive series of cases evaluated between January, 1999 and July, 2001. All clients that attended the practice were entered into the database. The data was provided to one of the authors (RB) in an anonymised form for analysis. Approval for this approach was granted by the Faculty Human Ethics Committee of La Trobe University, Bundoora, Australia. The sample of 93 non-TBI participants (61 male, 30 female) ranged in age from 18 to 49 years (mean = 34.16, *SD* = 8.5) and in education from 9 to 22 years (mean = 12.5, *SD* = 2.68).

The non-TBI participants were recruited as part of a normative research project, and were drawn from among community volunteers in south-east Queensland. All non-TBI participants were screened for neurological or psychiatric impairment. Refer to Table 1 for participant characteristics.

Measures

All of the TBI participants underwent comprehensive neuropsychological assessment which included the Wechsler Adult Intelligence Scale — — Third Edition (WAIS-III), the Wechsler Memory Scale — Third Edition (WMS-III), the Boston Naming Test (BNT), the FAS test, the Rey 15 item test, the Stroop Colour Word test and the Trail Making Test (TMT) parts A and B. The non-brain injured participants were assessed on the matrix reasoning, block design, arithmetic, digit span and letter number sequencing subtests from the WAIS-III.

Procedure

While the developers of the WAIS-III (Wechsler, 1997) have suggested that there are four problem types within the subtest (pattern completion, classification, analogy and serial reasoning), they did not provide a definition of the category types or identify which items were of which category type in the published descriptions of the test. We thus contacted Psychological Corporation and requested information regarding category definitions and the allocation of items to category type. Psychological Corporation unfortunately did not provide their category definitions; however, they did provide their item allocations and the method they used to arrive at these allocations. The information provided indicated that their item allocation was determined through group discussion as to which was the most likely used method of solving the item. Therefore it was possible that items could fall into more than one category.

Due to the lack of category definitions as well as the possibility that the items could fall into more than one category, we undertook a process to assign items to category by empirical means. This process began with the generation of category descriptions. The following definitions were thus developed:

TABLE 1

Demographic Details of the TBI and Control Samples

Variable	Mean	*SD*	Min–Max
Age (years)			
TBI	37.33	9.62	17–58
Control	34.16	8.50	18–49
Education (years)			
TBI	11.23	2.161	7–18
Control	12.50	2.68	9–22
TPI (months) TBI	60.35	62.00	6–418
TBI (days) PTA	27.42	62.11	.06–418

Note: All variables non significant at $p < .05$.

- pattern completion: to make whole or entire based upon a pattern
- classification: a systematic arrangement in groups according to an established criterion
- analogy: inference from a parallel case (i.e., resemblance in some particulars)
- serial reasoning: drawing conclusions from facts or premises arranged in an ordered sequence.

Three clinical neuropsychologists (two second year doctoral candidates and a practitioner with more than 15 years experience) were asked to classify the items into item type using the provided definitions.

The item allocations provided by Psychological Corporation were entered alongside our own rater determined allocations to determine the correlation between our own and the Psychological Corporation categories.

Interrater reliability correlations were conducted to determine the reliability of the established categories. These are presented in Table 2.

RESULTS

Prior to the formal statistical analyses, all variables were examined for accuracy of data entry and missing values. Missing data for the education, time post-injury and posttraumatic amnesia variables were replaced by the mean of all cases for that diagnostic group, but this only applied in less than 2% of the sample.

Category Allocations and Interrater Reliability

For our own research defined categories, the items were assigned as follows:

- pattern completion: 4, 5, 7, 13, 15, 16, 18
- classification: 1, 2
- analogy: 3, 6, 8, 10, 11, 14, 17, 19, 20, 21, 23, 25, 26
- serial reasoning: 9,12, 22, 24

The Psychological Corporation defined categories were assigned as follows:

- continuous and discrete pattern completion: 2, 4, 5, 6, 13, 15,16, 17, 18
- classification: 26
- analogy reasoning: 1, 3, 7, 8, 10, 11, 14, 19, 21, 23
- serial reasoning: 9, 12, 20, 22, 24, 25.

The interrater reliability for our categories was high with all correlations above .8 and statistically significant ($p < .0005$). The Psychological Corporation were also analysed by the correlational analysis, and indicated a somewhat lower correlation with the research defined categories at about .7. Refer to Table 2 for the interrater reliability correlations.

A one-way ANOVA was conducted on the TBI and the non-TBI participants to determine whether any differences were present between the overall MR performance of the groups and specifically how each performed on both the Psychological Corporation and the determined item categories. The test statistics are presented in Table 3.

Both the TBI and control groups performed differentially across the overall score and on each of the item categories, except for the classification items. Using the Psychological Corporation defined categories the TBI and control group performed significantly differently on all categories.

A repeated measures Analysis of Variance (ANOVA) with post hoc comparisons was then conducted to determine how both the TBI and non-TBI participants performed across the item categories as determined by the Psychological Corporation and ourselves. As each individual did not complete the full matrix reasoning subtest (each participants completed a different number of items depending on when they reached the discontinue criteria), the scores were rendered as a number of items correct for each item category up to the discontinue rule and were then treated as a

TABLE 2

Interrater Reliability of the Assignment of Items to Category for the MRT

Rater	1	2	3	Psychcorp
1	1.00	0.862	0.862	0.712
2	0.862	1.00	0.886	0.726
3	0.862	0.886	1.00	0.726
Psychcorp	0.712	0.726	0.726	1.00

TABLE 3
One Way Comparisons Between the TBI and Control Samples (Mean [*SD*]) Comparing the Research Defined (RD) Categories of Item Type and the Psychological Corporation (PC) Categories of Item Type

Variable	TBI	Control	*F*(1, 164)	Sig.
MR scaled score	8.3 (3.2)	12.4 (5.5)	25.45	0.000*
Analogy (RD)	55.2 (2.6)	71.2 (91.8)	25.18	0.000*
Classification (RD)	97.9 (1.2)	100.0 (0)	3.92	0.050*
Pattern completion (RD)	71.2 (3.2)	85.1 (1.7)	13.98	0.000*
Serial reasoning (RD)	43.2 (4.1)	66.1 (3.9)	11.51	0.001*
Analogy (PC)	68.6 (2.1)	81.8 (2.1)	26.33	0.000*
Classification (PC)	2.8 (2.0)	11.1 (3.7)	3.09	0.081
Pattern completion (PC)	73.0 (2.8)	85.1 (1.7)	15.01	0.000*
Serial reasoning (PC)	41.3 (4.1)	59.4 (3.4)	16.56	0.000*

percentage correct of the number of items in that category up to the point of discontinuation (i.e., if participant A was administered five items out of Category A that contains 10 items, but only answered correctly four of the five items administered, then the score of four out of five items, 80%, would be used).

A repeated measures ANOVA indicated that both the TBI and control group performed differentially across the item categories as defined by both ourselves and the Psychological Corporation. Mauchly's Test of Sphericity was significant for both the TBI and control group analyses, indicating that the assumption of sphericity was violated and therefore the Greenhouse-Geisser correction was employed. For our own defined categories, repeated measures ANOVA was significant, TBI: F (2.034, 144.428) = 89.904, $p < .0005$, Control: F (1.643, 116.622) = 48.178, $p < .0005$, indicating that there was a difference in performance of both groups across the categories.

Post hoc analyses (Student Newman–Keuls: SNK), using the Bonferroni adjustment for multiple comparisons, indicated that performance across all category types was significantly different for the TBI group. For the control group there was no significant difference between performance on the analogy and serial reasoning categories (refer to Table 4).

For the Psychological Corporation defined categories, repeated measures ANOVA was significant, TBI: F(2.231, 158.397) = 151.041, $p < .0005$, Control: F(2.294, 162.902) = 172.735, $p < .0005$, indicating that there was a difference in performance of both groups across the categories. Post hoc analyses (SNK), using Bonferroni adjustment for multiple comparisons, indicated that there was a significant difference between all categories except for performance on the analogy and pattern completion categories for both the TBI and control group. Refer to Table 5.

A regression analysis was then undertaken to investigate the differential influence of each of the item category types to the final matrix reasoning scaled score.

Regression Analyses (Psychological Corporation Categories)

TBI Group

Age, level of education and length of post-traumatic amnesia were not significantly correlated with the matrix reasoning scale score. Time post-injury was negatively correlated with MR scaled score, $r = -0.24$, $p = .02$. MR scaled score was positively correlated with all the item categories (analogy: r = 0.63, $p < .0005$; classification: $r = .27$, $p = .012$; pattern completion: $r = .47$, $p < .0005$; serial reasoning: $r = .60$, $p < .0005$).

A multiple regression analysis using MR scaled score as the outcome variable and time post-injury and analogy, classification, pattern completion and serial reasoning categories as the predictor variables was then undertaken. In this initial equation, analogy, pattern completion and serial reasoning categories were significant predictor variables. Time post-injury did not reach significance and the classification category approached significance. The regression was recomputed without time post-injury as a predictor variable. In this second equation, classification was still a non-significant predictor and thus it was removed and the regression analysis was recomputed. The final regression equation indicated that 69.2% of the variance of the MR scaled score was explained by performance on the analogy, pattern completion and serial reasoning categories. ANOVA

TABLE 4

Means (*SD*) of the TBI and Control Group for Our Research Defined Categories of Item (Factor I and J Represent Univariate Comparisons Between Each Item Type; i.e., 1 vs. 2)

Factor	Mean	Std. Error	Factor (I)–	Factor (J)	Mean diff. (I–J)	Std. Error	Sig.
TBI							
1 Analogy	55.2	2.6	1	2	–42.736	2.275	0.000
				3	–16.000	3.377	0.000
				4	12.00	3.510	0.006
2 Classification	97.9	1.2	2	3	26.736	2.693	0.000
				4	54.736	3.974	
3 Pattern completion	71.2	3.2	3	4	28.000	4.776	0.000
4 Serial reasoning	43.2	4.1					
Control							
1 Analogy	71.5	1.8	1	2	–28.514	1.758	0.000
				3	–13.639	2.451	0.000
				4	5.361	3.443	0.743
2 Classification	100	0.0	2	3	14.875	1.711	0.000
				4	33.875	3.857	0.000
3 Pattern completion	85.1	1.7	3	4	19.000	4.299	0.000
4 Serial reasoning	66.1	3.9					

TABLE 5

Means (*SD*) of the TBI and Control Group for the Psychological Corporation Defined Categories of Item (Factor I and J Represent univariate Comparisons Between Each Item Type; i.e., 1 vs. 2)

Factor	Mean	Std. Error	Factor (I)–	Factor (J)	Mean diff. (I–J)	Std. Error	Sig.
TBI							
1 Analogy	68.6	2.1	1	2	65.806	2.688	0.000
				3	–4.444	3.424	1.000
				4	27.250	3.465	0.000
2 Classification	2.8	2.0	2	3	–70.250	3.196	0.000
				4	–38.556	4.288	0.000
3 Pattern completion	73.0	2. 8	3	4	31.694	4.830	0.000
4 Serial reasoning	41.3	4.1					
Control							
1 Analogy	81.8	2.1	1	2	70.708	3.709	0.000
				3	–3.306	2.240	0.867
				4	22.458	3.300	0.000
2 Classification	11.1	3.7	2	3	–74.014	3.972	0.000
				4	–48.250	4.757	0.000
3 Pattern completion	85.1	1.7	3	4	25.764	3.590	0.000
4 Serial reasoning	59.4	3.4					

indicated that this was significant, F (3, 68) = 50.88, $p < .0005$.

The correlation matrix was examined for multicollinearity and singularity of the variables in the regression equation. All correlations amongst the category variables were less than 0.70 indicating the absence of multicollinearity and singularity and all variables were subsequently entered into the regression equations. In conducting the regression equation on the control group using the research defined categories, the classification category variable was not entered into the regression equation as it was found to be a constant value, as all the participants answered these items correctly.

Control Group

Initial analysis of correlations indicate that age, education and analogy, classification and serial reasoning categories were not significantly correlated with MR scaled score. Only the pattern completion category was correlated with MR scaled score, $r = 0.21$, $p = .041$.

Multiple regression analysis using the enter method indicated that 6.5% of the variance of the MR scaled score was explained by analogy, classification, pattern completion and serial reasoning categories. The regression equation indicated that none of these variables was significant. ANOVA was also non-significant, $F(4, 67) = 1.16$, $p = .339$.

Regression Analyses (Research Defined Categories)

TBI Group

Age, level of education and length of post-traumatic amnesia were not significantly correlated with the matrix reasoning scaled score. Time post-injury was negatively correlated with MR scaled score, $r = –0.24$, $p < .02$. MR scaled score was positively correlated with all the item categories (analogy; $r = 0.74$, $p < .0005$; classification: $r = 0.41$, $p < .012$; pattern completion: $r = 0.48$, p < .0005; and serial reasoning: $r = 0.58$, $p < .0005$).

An enter method multiple regression analysis using MR scaled score as the outcome variable and time post-injury and the analogy, classification, pattern completion and serial reasoning categories as the predictor variables was then performed. In this initial equation neither time post-injury nor the classification category approached significance. Therefore the regression analysis was recomputed with these variables removed. The final regression equation indicated that 65.5% of the variance of the MR scaled score was explained by performance on the analogy, pattern completion and serial reasoning categories. ANOVA indicated that this was significant, $F(3, 68) = 43.02$, $p < .0005$.

Control Group

Age, education and the analogy and serial reasoning categories were not significantly correlated with MR scale score. Only the pattern completion category was correlated with MR scaled score, $r = 0.21$, $p = .041$. The classification category was successfully completed by all participants and was therefore a constant in the analysis and no correlation was computed nor was it entered into the regression analyses.

Multiple regression analysis using the enter method indicated that 12% of the variance of the MR scaled score was explained by analogy, pattern completion and serial reasoning categories. The regression equation indicated that pattern completion was the only significant predictor in the equation, $p < .006$. ANOVA indicated that this was significant, $F(3, 68) = 3.105$, $p = .032$.

DISCUSSION

Qualitative analysis of test instruments has been a long tradition in neuropsychological assessment. This study extended this type of analysis to the Matrix Reasoning (MR) subtest of the WAIS-III. The first task undertaken in the study was to define the item types and to assign each item to a type. Interrater reliability as determined using our research defined criteria indicated a significant pattern of correlation consistently above 0.8. The pattern of assignment as determined by Psychological Corporation yielded a somewhat lower pattern of interrater reliability.

Comparison of the item types between the TBI and control group using both our research defined categories as well as the Psychological Corporation categories indicated that for all category types the performances were significantly different for the TBI and the non-injured group. In the control group there were no differences between performance on the analogy and serial reasoning categories. Determination of the level of performance between the control and TBI groups on the various categories revealed that the TBI group differed significantly across all categories except for the classification group. With the Psychological Corporation categories, the TBI and control groups differed significantly on all categories.

Multiple regression analysis revealed that only a small proportion of the variance in matrix reasoning performance for the control group was explained by the analogy, classification, pattern completion and serial reasoning categories. This outcome probably occurred as a result of the restricted range of test scores in the control group.

In the case of the TBI group however, 65.5% of the variance of the matrix reasoning scale score was explained by performance on analogy, pattern completion and serial reasoning.

The results of this investigation indicate that qualitative analysis of the performance of TBI participants can be usefully undertaken with item types on the matrix reasoning subtest. This observation adds support to the use of the qualitative of aspects of the performance to characterise the nature of the functions assessed by this instrument. It is also interesting to note that the nature of the explained variance in the regression equations differed markedly as a function of previous history of TBI, with the TBI patients having much higher amounts of the variance explained than for the non-impaired controls. This observation supports the notion that the item type analysis is responsive to injury status.

There were a number of weaknesses identified in the current investigation which may undermine the findings to some degree. The nature of assignment of items to type remains something that has to be determined on the basis of face validity rather than by any more robust statistical technique. Nonetheless, it was possible to get reasonably good concordance between rates for assignment of items to category. The investigation was also undermined to some degree, by the fact that the TBI group did not complete all items on the scale. Nonetheless in the clinical practice situation this would indeed be the manner in which such data would be gathered.

A second and more concerning aspect of the study was the number of items included in the classification group of items. In our research determined groups this item type is made up of only two items (items 1 and 3) which would not ordinarily be given in a standard administration of the test (which begins with item 4). As such, if the administration began at item 4, the only time in which items 1 and 3 would be administered would be in retrograde order if either items 4 or 5 were failed. While the use of this category could bias the data collection in our study, it did not do so because in each case this item type was excluded from the multiple regression analysis due to lack of variability.

A second problem pertains to the use of the Psychological Corporation Category of classification which contains only a single item, item number 26. This item as the last and most difficult item on the task would be unlikely to be reached by any but the most high functioning of participants. As a consequence, the interpretation of the performance on the classification category in both schemes must be viewed with the utmost caution in any qualitative interpretation of the performance. Clearly, considerably more work needs to be done in this area regarding what categories of item do exist in the task and this study is merely the first step in trying to define these.

A third concern with the category assignment is that some categories may inadvertently contain more difficult items than others. While we have attempted to overcome this problem by the use of the percentage of items correct up to the discontinue rule, some problems with this solution persist. For example, some categories (e.g., the Psychological Corporation serial reasoning category consists of more difficult items — four of the six items are in the latter part of the subtest; i.e., 20, 22, 24 and 25). As a result, there are likely to be fewer participants contributing to the scores in these categories, thus the scores observed may be reflecting item difficulty per se rather than any specific aspect of cognitive processing.

A number of possible avenues for further investigation arise from the current study. It would certainly be worthwhile to determine how much of the variance in MR performance is explained by the constructs of processing speed and working memory in line with the previous work of Salthouse (1993). Both of these constructs have been demonstrated to be responsive to the effects of TBI at various levels of severity (Psychological Corporation, 1997).

Preliminary analysis using the TBI sample employed in this study indicates that this will be a fruitful area of enquiry as both of these constructs correlate strongly (all *p*s $< .01$) with the item types discussed in this article (processing speed with pattern completion, $r = .31$; classification, $r = .29$; analogy, $r = .62$; serial reasoning, $r = .41$; working memory with pattern completion, $r = .32$; classification, $r = .29$; analogy, $r = .65$; serial reasoning, $r = .51$). This presents the possibility of a fruitful approach for further study of these issues.

The use of other groups of brain impaired participants may also be worthwhile in the further investigation of this effect and to determine if there is specificity by item category in these groups. It would also be worthwhile to specify the cognitive operations that are necessary to conduct each item type by a componential analysis of task performance (e.g., Crowe, 1998).

The results of the current investigation indicate that there are differential patterns of performance by item type in a group of traumatic brain injured individuals in comparison to non-injured controls. The classification type problems are the most easy (and with the fewest items), while analogy and serial reasoning problems are the most difficult. The study provides a further validation and rationale for the notion of qualitative analysis of performance on neuropsychological instruments in the assessment situation which can act as a basis for further specification of these constructs in future investigations.

REFERENCES

Babcock, R.L. (1994). Analysis of adult age differences on the Raven's Advanced Progressive Matrices Test. *Psychology and Aging, 9(2),* 303–314.

Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing of the Raven Progressive Matrices Test. *Psychological Review, 97(3),* 404–431.

Crowe, S.F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search and motor speed to performance on Parts A and B of the Trail Making Test. *Journal of Clinical Psychology, 54,* 585–591.

Dikmen, S.S., Machamer, J., Winn, H.R., & Temkin, N. (1995). Neuropsychological outcome at 1-year post head injury. *Neuropsychology, 9,* 80–90.

Donders, J., Tulsky, D.S., & Zhu, J. (2001). Criterion validity of new WAIS-III subtest scores after traumatic brain injury. *Journal of the International Neuropsychological Society, 7,* 892–898.

Dugbartey, A.T., Sanchez, P.N., Rosenbaum, J.G., Mahurin, R.K., Davis, J.M., & Townes, B.D. (1999). WAIS-III Matrix Reasoning performance in a mixed clinical sample. *The Clinical Neuropsychologist, 13(4),* 396–404.

Green, K.E., & Kluever, R.C. (1991). Structural properties of Raven's Coloured Progressive Matrices for a sample of gifted children. *Perceptual and Motor Skills, 72,* 59–64.

Gutman, R. (1974). Genetic analysis of analytical spatial ability: Raven's Progressive Matrices. *Behavior Genetics, 4(3),* 273–284.

Marshalek, B., Lohman, D.F., & Snow, R.E. (1983). The complexity continuum in the radix and hierarchical models of intelligence. *Intelligence, 1983, 7(2),* 107–127.

Psychological Corporation (1997). *Technical manual of the WAIS-III and the WMS-III.* San Antonio: Harcourt Brace and Company.

Raven, J.C. (1948). The comparative assessment of intellectual ability. *British Journal of Psychology, 39,* 92–99.

Salthouse, T.A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology, 84,* 171–199.

Soukup, V.M., Harrell, E,H., & Clark, T. (1994). Right hemispace presentation and left cueing on Raven's Coloured Progressive Matrices among right brain-damaged neglect patients. *Brain Injury, 8(5),* 449–455.

Tulsky, D.S., Saklofske, D.H., & Zhu, J. (2003). Revising a standard: An evaluation of the origin and development of the WAIS-III. In D.S. Tulsky, D.H. Saklofske, G.J. Chelune, R.K. Heaton, R.J. Ivnik, et al. (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 43–92). San Diego, CA: Academic Press.

Villardita, C. (1985). Raven's Coloured Progressive Matrices and intellectual impairment in patients with focal brain damage. *Cortex, 21,* 627–634.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale — III (WAIS-III) manual.* New York: The Psychological Corporation.