A comparison of single and multi-test working memory assessments in predicting academic achievement in children

Kerry A. Chalmers and Emily E. Freeman

School of Psychology

University of Newcastle, Australia

Corresponding Author:

Kerry A. Chalmers

School of Psychology

University of Newcastle, University Drive, Callaghan, NSW 2308, Australia

Email: kerry.chalmers@newcastle.edu.au

Abstract

Children assessed as having low working memory capacity have also been shown to perform more poorly than their same-aged peers in measures of academic achievement. Early detection of working memory problems is therefore an important first step in reducing the impact of a working memory deficit on the development of academic skills. In this study, we compared a single-test assessment, the Working Memory Power Test for Children (WMPT) and a multi-test assessment, the Automated Working Memory Assessment (AWMA), in their ability to predict academic achievement in reading, numeracy and spelling. 132 Australian school children (mean age 9 years, 9 months) participated in the research. Strong positive correlations between the WMPT and AWMA total scores were found, indicating good convergent validity of the single and multi-test measures. WMPT scores correlated with each of the four AWMA subtests designed to assess verbal and visuospatial short-term and working memory. WMPT and AWMA scores separately predicted performance on Word Reading, Numerical Operations and Spelling. Compared with either measure alone, the WMPT and the AWMA in combination predicted more of the variance in Word Reading and Numerical Operations, but not in Spelling. Theoretical and practical implications of these findings are discussed.

KEYWORDS: working memory; assessment; academic achievement; literacy; numeracy

Working memory and academic achievement in children: A comparison of single and multi-test working memory assessment methods

Working memory has been referred to as the ability to briefly store and manipulate information that is needed to perform complex tasks such as language comprehension, learning and reasoning (Baddeley, 1992; 2000). Children with a deficit in working memory have been found to have a lower level of academic achievement in literacy and numeracy (Alloway, Gathercole, Kirkwood, & Elliott, 2009). The extent to which working memory predicts academic achievement has been shown to depend on the type of working memory task used. For example, some authors have reported a relationship between performance on a verbal working memory task and literacy skills (e.g., Niedo, Abbott, & Berninger, 2014). Others have found a relationship between performance of visuospatial working memory tasks and mathematical ability (e.g., Bull, Espy, & Wiebe, 2008; Dumontheil & Klingberg, 2012; McLean & Hitch, 1999). In the current study, we compare two forms of working memory assessment, a single-test measure, the Working Memory Power Test for Children (WMPT), and a multi-test battery, the Automated Working Memory Assessment (AWMA), in their ability to predict performance in reading, numeracy and spelling.

Models of Working Memory

The most commonly described model of working memory is the multi-component model developed by Baddeley and his colleagues (Baddeley & Hitch, 1974). This model assumes that incoming information falls into one of two broad categories, verbal or visuospatial. According to the model, verbal and visuospatial information is temporarily held by the phonological loop and visuospatial sketchpad, respectively. The phonological loop and visuospatial sketchpad are thought of as slave systems that are controlled by the central executive, which is responsible for directing attention and controlling resources. The more recent versions of the model have also included a fourth component, the episodic buffer (Baddeley, 2000, 2007). The episodic buffer is assumed to have links to long-term memory and assists with integrating information into a single unit/episode as well as reinstating the chronological ordering (time sequencing) of information, such as the storyline in a novel.

In addition to the multi-component model, a number of less complex models of working memory have been proposed. One of the simplest is Cowan's (1999) embedded process model, which views working memory as a limited capacity attentional focus that is directed towards that currently active area of long-term memory. While capacity and control of focus appear to rely on different brain regions, there is no differentiation between specific systems required for processing verbal and visuospatial information in this model (Cowan, 2005). Any mechanism that contributes to the desired outcome is regarded as participating in the working memory system.

Working Memory and Academic Achievement

The relationship between working memory and academic achievement during childhood has been examined in a number of studies (e.g., Alloway & Passolunghi, 2011; Cameron, Glyde, & Dillon, 2014; Metcalfe, Ashkenazi, Rosenberg-Lee & Menon, 2013; Niedo et al., 2014; St Clair Thompson & Gathercole, 2006). Some studies have examined working memory in children from the perspective of the multi-component model. Often these studies find children with reading difficulties have lower scores on all of the components of working memory postulated by this model. For example, Siegel and Ryan (1989) showed children with reading, but not numeracy difficulties, scored significantly lower than their peers on both verbal and visuospatial working memory tasks. More recently, Swanson, Kehler and Jerman (2010) found children with reading difficulties had lower scores on both verbal and visuospatial working memory tasks, but no difference in strategy knowledge compared with their peers.

While some researchers have concluded that both verbal and visuospatial working memory components are related to numeracy (e.g., Zheng, Swanson & Marcoulides, 2011), others have found that children experiencing problems with arithmetic may be impaired in some, but not all aspects of working memory. For example, Bull et al. (2008) found mathematical achievement at the end of Year 3 is largely predicted by visuospatial working memory. More recently, Szucs, Devine, Soltesz, Nobes and Gabriel (2013) found that impaired visuospatial short-term and working memory, but not verbal short-term or verbal working memory, were related to developmental dyscalculia.

McLean and Hitch (1999) suggested that a deficit in visuospatial processing may underlie difficulties in reading and/or writing numbers. Mammarella, Lucangeli and Cornoldi (2010) assessed visual and spatial processing separately, and concluded that spatial working memory rather than visual processing was related to difficulties in both number ordering and calculation. Passolunghi and Mammarella (2010) also found a deficit in spatial rather than visual processing was related to poor problem-solving ability.

Deficits in working memory can also lead to difficulties in forming long-term memory representations of basic arithmetic facts, presumably due to information decaying too quickly to allow the formation of the relevant associations. Children with limited working memory capacity may also have difficulty retrieving arithmetic facts from long-term memory for timely use in solving an arithmetic problem (McLean & Hitch, 1999).

Given the established links between working memory and academic achievement, it is important that children who are struggling academically be assessed for a possible deficit in working memory so that appropriate interventions can be implemented. The availability of easy-to-administer assessments of working memory has the potential to facilitate this process.

Assessment of Working Memory in Children

One of the most commonly used test batteries for the assessment of working memory in children (and adults) is the Automated Working Memory Assessment (Alloway, 2007). The AWMA is designed to assess each of the components assumed in Baddeley's multi-component model. The AWMA short-form consists of four subtests designed to assess verbal short-term memory (STM), visuospatial STM, verbal working memory and visuospatial working memory.

The WMPT has been specifically designed to assess working memory performance in children. In contrast to the AWMA, the WMPT is compatible with a unitary view of working memory such as assumed in Cowan's (1999) embedded process model. The WMPT uses non-verbal stimuli (i.e., line drawings of animals familiar to children) and simple English instructions (e.g., Swap 1 and 2), making it suitable for children of different language abilities. An advantage of the test over existing batteries such as the AWMA is that it is fully automated. The examiner is not required to enter the participant's responses to each question. Once the child is logged in to the test, it can be completed with minimal supervision.

The Current Study

There were two main aims of the current study. The first was to examine the convergent validity of the WMPT, a single-test measure, against the AWMA, a multi-test assessment of working memory. Convergent validity was assessed for each of the four components of working memory assessed by the AWMA (verbal STM, visuospatial STM, verbal working memory and visuospatial working memory) as well as for the AWMA total score. It was predicted there would be a positive correlation between WMPT and AWMA total scores, indicating convergent validity. The extent to which the WMPT assesses verbal versus visuospatial components of working memory could be indicated by the strength of the relationship between WMPT scores and each of the AWMA subtests. While the stimuli used in the WMPT are visually presented images, making it possible to perform the task by visualising the animals "changing places", we expected participants would also use verbal coding (e.g., "swap the pig with the cat") when performing the tasks. We therefore expected the WMPT could be related to both the verbal and visuospatial short-term memory subtests of the AWMA (i.e., Digit Span and Dot Matrix). Because the WMPT has a high processing requirement (holding and manipulating information on each swap trial) we also predicted a strong correlation between performance on the WMPT and the verbal and visuospatial working memory subtests of the AWMA (i.e., Listening Recall and Spatial Recall).

Our second aim was to examine the relative contribution of performance on the WMPT and the AWMA to the prediction of academic achievement in reading, numeracy and spelling as assessed by the Wechsler Individual Achievement Test – Second Edition, Australian Abbreviated (WIAT-II, Wechsler, 2007). We expected both the AWMA and the WMPT measures of working memory would predict academic achievement in reading, numeracy and spelling. There were two main questions: First, whether the proportion of variance in academic achievement accounted for by each working memory measure (i.e., WMPT and AWMA) would be similar; and second, whether academic achievement is better predicted by performance on both working memory assessments in combination compared with either alone.

Method

Participants

One hundred and thirty-two children (Males = 66, Females = 66) attending primary school in regional NSW, Australia, participated in the research.[1] The mean age of the children was 9 years, 9 months (ranging from 8 years, 8 months and 11 years, 1 month). All children spoke fluent English. Ethics approval for this research was granted by the Human Research Ethics Committee of the University of Newcastle and the Catholic Schools Office, Diocese of Maitland-Newcastle.

Measures

Working Memory Power Test for Children (WMPT, Lewis Cadman Consulting Pty Ltd, Sydney, Australia).[2]The WMPT is an online assessment used to measure children's working memory performance, and is designed so that children can complete it independently or under supervision. It assesses both confidence and accuracy in performance.[3] The WMPT has five levels of increasing difficulty (0-Swap, 1-Swap, 2-Swap, 3-Swap, 4-Swap). In the 0-Swap condition, the child is presented with a display consisting of three items (e.g., *pig, cat, duck*) numbered 1, 2, 3, from left to right (see Figure 1 for an example of the display). The child is asked to remember the items and their position in the display. They are then shown an answer screen in which the same three items are displayed in each possible order and are asked to select the option (by mouse click) that represents the correct order of the items. For the 0-Swap level, the correct response is the sequence of items that is the same as the one shown at study (i.e., in our example, *pig, cat, duck*). Subsequent levels of the

WMPT involve mentally swapping the order of the items. In the 1-Swap condition (using the same example as shown in Figure 1), the instruction might be to "Swap 1 and 2". The child must mentally swap the order of the pictures and then select the option that represents the correct sequence of pictures after the swap has been made (i.e., *cat, pig, duck*, in our example). In the 2-Swap condition, the instruction might be to "Swap 1 and 2, then Swap 2 and 3" (answer: *cat, duck, pig*). The levels progress from 0 to 4 swaps, with task difficulty increasing as more consecutive swaps are required. After each test item, the child is asked to rate how confident they are that they answered the preceding question correctly using a 4-point scale (*I guessed the answer* through to *I definitely got it right*). There are five trials at each of the five levels of difficulty, for a total of 25 trials. Total correct (out of 25) is converted to a percentage. The WMPT has been shown to have good reliability (Cronbach's alpha =.85), calculated on scores from 170 Australian schoolchildren enrolled in Year 4 of primary school (Chalmers & Freeman, 2017). WMPT test scores have also been shown to predict academic achievement, suggesting good concurrent validity (Chalmers & Freeman, 2017).

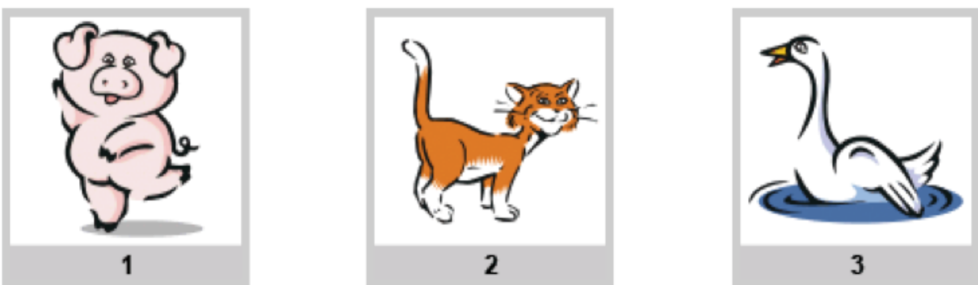

Figure 1. Example of a study trial presentation in the Working Memory Power Test for Children.

Automated Working Memory Assessment – Short Form (AWMA, Alloway, 2007). The AWMA is a well-researched measure of children's working memory performance. It is based on the multi-component model of working memory (Alloway, 2007). The short-form consists of four subtests, one for each verbal, visuospatial, short-term and working memory component. Verbal STM is measured using an auditory digit span task. Participants hear a string of digits and are asked to repeat the digits, in the correct order, back to the experimenter. Verbal working memory is measured using an auditory sentence recall task. Participants listen to a sentence and are asked to report whether or not the sentence is true or false (e.g., "Dogs have four legs"). After a series of sentences have been presented and classified as true or false, participants are asked to report the final word from each of the sentences in the order in which they heard them. Visuospatial STM is measured using a visually presented dot matrix task. Participants are presented with a 4x4 matrix. A series of red dots appears one at a time in different cells of the matrix. A blank matrix is then presented and the participant is required to point to each spatial location in which the red dot appeared, in the order they were presented. Visuospatial working memory is measured using a visuospatial recall task. Two symbols are presented on screen. The symbols are either identical, or one is a mirror image of the other. The symbol on the right is presented at 0, 120, or 240 degrees orientation and accompanied by a red dot. Participants are first asked to report whether the symbol on the right is the 'same' or 'different' to the symbol on the left. After a series of symbols have been presented, three black dots appear on the screen (at 0, 120, and 240 degrees orientation) and the participant is asked to point to the location in which each red dot had been presented, in the order they were presented. The AWMA is suitable for children and young adults aged between 4 and 22 years. Scores range from 0 to 54, 54, 36, and 42, for Digit Span, Dot Matrix, Listening Recall and Spatial Recall, respectively. The AWMA has good test-retest reliability (.89, .85, .88, and .79, for Digit Recall, Dot Matrix, Listening Recall and Spatial Recall subtests, respectively, Alloway, 2007).

Wechsler Individual Achievement Test - Second Edition, Australian Abbreviated (WIAT-II; Wechsler, 2007). The WIAT-II is a paper and pencil achievement test assessing Word Reading, Numerical Operations, and Spelling. Word Reading assesses letter identification skills and phonological awareness. Numerical Operations involves items that assess early maths skills (e.g., number recognition and counting) and higher calculation skills (e.g., solving equations). Spelling assesses the individual's ability to spell dictated letters, words and letter blends. The test is suitable for use in individuals from 5 to 85 years. Scores range from 0 to 131, 54, or 53, for Word Reading, Numerical Operations and Spelling, respectively. The WIAT-II is individually administered. Reliability for each subtest and the composite score is high, with reliability coefficients of .91 (Numerical Operations and Spelling), .96 (Word Reading) and .97 (Composite Score) for Australian children (Wechsler, 2007).

Procedure

Children participated in the research during their normal school day. Each child was tested individually, with the order of presentation of the WMPT, AWMA and WIAT-II counterbalanced across participants.[4] Total testing time was approximately 1½ hours, including breaks between each task.

Children completed the WMPT on a laptop computer. Written instructions were presented on the screen before each level (i.e., 0-Swap, 1-Swap etc.) commenced. The 0-Swap, 1-Swap and 2-Swap conditions started with two practice problems. All participants completed the levels in the same order (i.e., 0-Swap, 1-Swap, 2-Swap, 3-Swap, 4-Swap).

The AWMA was administered by a trained researcher. Each subtest was presented on a laptop computer using the AWMA computerised software. Children stated their responses to the experimenter, who then manually scored the response as correct or incorrect using the arrow keys on the keyboard. Children completed the subtests in the standard order, Digit Span, Dot Matrix, Listening Recall, and Spatial Recall. Each subtest began with a series of practice trials. For each subtest, the level of difficulty increased by one item (i.e., one extra digit, dot, sentence, or symbol). If the child successfully recalled four items at a given level, they progressed to the next level. If three mistakes were made within a level, the subtest ended. Raw and standardised scores for each participant for each subtest were obtained using the AWMA software.

The WIAT-II was administered by a trained researcher. Children completed the subtests in the standard order. The Word Reading subtest, where they were asked to read a list of words to the experimenter, was presented first. Next, in the Numerical Operations subtest, they were given a sheet of math problems to solve. For this subtest they were asked to show their working and were not given an eraser to correct mistakes. In the final subtest, Spelling, the experimenter read a list of words one at a time to the child, who then wrote the word on a standard response form. Children started each subtest at the Year 4 appropriate start point. For each subtest, if an error was made on any of the first three items, the preceding items were administered in reverse order until three consecutive correct scores were achieved. Subtests were discontinued if six (Numerical Operations and Spelling) or seven (Word Reading) consecutive errors were made. Each subtest was scored according to the WIAT-II manual.

Data Analysis

Descriptive statistics for both raw and standard scores on the AWMA and the WIAT-II are reported. Inferential statistics were performed using the raw scores. The dependent variable for the WMPT, total percent correct, was calculated as the number of correct responses out of 25 trials, converted to a percentage. Analysis of variance (ANOVA) was used to examine whether the order of presentation of the WMPT, AWMA and WIAT tests affected performance in these tasks. Pearson correlations were used to examine the convergent validity between the two working memory tests (i.e., WMPT and AWMA) and the relationship between working memory and academic achievement (as assessed by the WIAT-II). Hierarchical regression analyses were conducted to examine whether the contribution of each of the working memory measures to the prediction of academic achievement changed after performance on the alternative measure of working memory was controlled.

Results

Preliminary analysis showed there was no effect of counterbalancing order on performance in the WMPT, AMWA, or WIAT-II. Mean accuracy for the WMPT was 61% (*SE* = 1.7%). Mean accuracy (raw and standard scores) for each AWMA subtest, and overall, and for the WIAT-II subtests are presented in Table 1. The standard scores for the AWMA and WIAT-II subtests indicate that average performance was within the normal range (i.e., the sample mean is within one standard deviation of the population mean, where *M* = 100 and *SD* = 15).

Table 1

Mean Raw and Standard Scores (Standard Error in Parentheses) for the Automated Working Memory Assessment (AWMA) and Wechsler Independent Achievement Test (WIAT-II)

	Raw score	**Standard score**
AWMA		
Digit Span	28.62 (0.38)	103.25 (1.43)
Dot Matrix	22.86 (0.44)	102.39 (1.47)
Listening Recall	12.74 (0.29)	103.80 (1.22)
Spatial Recall	19.30 (0.46)	108.47 (1.19)
Total	83.52 (1.14)	104.12 (1.22)
WIAT-II		
Word Reading	102.70 (0.75)	99.65 (0.86)
Numerical Operations	20.19 (0.33)	95.88 (1.21)
Spelling	31.86 (0.57)	102.64 (1.33)

Pearson Correlation Analysis of the Relationship Between WMPT and AWMA Scores

Pearson correlations between WMPT and AWMA scores are presented in Table 2. WMPT scores were positively correlated with all four AWMA subtests and AWMA total. The strong correlation between WMPT accuracy and the AWMA total score suggests good convergent validity between the two forms of working memory assessment. A weak but significant correlation was found between WMPT and Digit Span, which is assumed to assess verbal STM. Moderate correlations were found between WMPT and Dot Matrix, Listening Recall and Spatial Recall, which are assumed to assess visuospatial STM, verbal working memory, and visuospatial working memory, respectively.

Table 2

Pearson Correlations Examining Relationships Between Performance on the Working Memory Power Test for Children (WMPT), the Automated Working Memory Assessment (AWMA), and Achievement in Reading, Numeracy and Spelling (as assessed by the Wechsler Individual Achievement Test)

	2	3	4	5	6	7	8	9
1. WMPT	.503***	.260**	.415***	.420***	.377***	.339***	.580***	.283**
2. AWMA total	-	.620***	.810***	.622***	.814***	.358***	.509***	.304***
3. Digit Span		-	.297**	.232**	.289**	.391***	.286**	.259**
4. Dot Matrix			-	.367***	.583***	.157	.397***	.159
5. Listening Recall				-	.372***	.324**	.248**	.267**
6. Spatial Recall					-	.214*	.496***	.222*
7.Word Reading						-	.431***	.727***
8. Numerical Operations							-	.522***
9. Spelling								-

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

Relationship Between Working Memory and Academic Achievement

Pearson correlations between the working memory test scores and Word Reading, Numerical Operations and Spelling scores (as assessed by the WIAT-II) are also presented in Table 2. WMPT scores were positively correlated with each measure of academic achievement. The correlations between WMPT scores and Word Reading, Numerical Operations and Spelling were moderate, strong, and weak, respectively.

AWMA total scores were moderately correlated with Word Reading and Spelling, and strongly correlated with Numerical Operations. Of the AWMA subtest scores, Digit Span and Listening Recall were moderately correlated with Word Reading, and weakly correlated with Numerical Operations and Spelling. Digit span is assumed to assess verbal STM and Listening Recall is assumed to assess verbal working memory, suggesting these verbal tasks have a stronger relationship with reading than either numeracy or spelling. AWMA Dot Matrix and AWMA Spatial Recall were moderately correlated with Numerical Operations, suggesting spatial STM and spatial working memory are both related to numeracy. AWMA Spatial Recall was weakly correlated with Word Reading and Spelling, suggesting a contribution of spatial working memory to these primarily verbal tasks.

Working Memory as a Predictor of Academic Achievement

A series of six hierarchical multiple regressions were conducted to examine whether the WMPT and the AWMA together predicted academic achievement over and above either of these measures alone. In each model, scores for the WMPT and each of the four AWMA subtests, as well as gender and age, were the predictors. Academic achievement in Word Reading, Numerical Operations or Spelling was the criterion variable. Two models were computed for each criterion variable, one in which the WMPT score (and age and gender) was entered at Step 1, with the AWMA subtest scores entered in Step 2, and one in which the AWMA subtest scores (and age and gender) were entered in Step 1, with the WMPT score entered in Step 2. A summary of the results is presented in Table 3.

Table 3

Results from Hierarchical Regression Analyses Examining the Relative Contribution of Working Memory as Assessed by the Working Memory Power Test for Children (WMPT) and the Automated Working Memory Assessment (AWMA) to the Prediction of Academic Achievement in Word Reading, Numerical Operations and Spelling as Assessed by the Wechsler Individual Achievement Tests (WIAT-II)

	Predictors	R^2	F	β	*t*
Word Reading					
	(a) Step 1 – WMPT entered	.12	5.65**		
	Gender			.04	<1
	Age			-.01	<1
	WMPT			.34	4.03***
	(b) Step 1 – AWMA entered	.22	5.81***		
	Gender			.05	<1
	Age			-.02	<1
	Digit Span			.33	3.84***
	Dot Matrix			-.07	<1
	Listening Recall			.25	2.88**
	Spatial Recall			.07	<1
	(c) Step 2 – WMPT & AWMA	.25	5.85***		
	Gender			.04	<1
	Age			<-.01	<1
	WMPT			.21	2.24*

When the WMPT score was entered at Step 1, significant models were produced for Word Reading, R^2 = .12, F(3,128) = 5.65, p = .001, f^2 = .14, Numerical Operations, R^2 = .36, F(3,128) = 23.89, p < .001, f^2 = .56, and Spelling, R^2 = .08, F(3,128) = 3.83, p =.011, f^2 = .09, explaining between 8% and 36% of the variance in achievement scores, with effect sizes (using Cohen's, 1992, conventions for f^2) ranging from small to large. Adding the four AWMA subtest scores in Step 2, produced a significant R^2 change in both Word Reading, ΔR^2 = .13, F(4,124) = 5.42, p < .001, and Numerical Operations, ΔR^2 =.10, F(4,124) = 6.00, p < .001, but not in Spelling, ΔR^2 =.06, F(4,124) = 2.26, p = .066.

Similarly, when the four AWMA subtest scores were entered in Step 1, significant models for Word Reading, R^2 = .22, F(6,125) = 5.81, p < .001, f^2 = .28, Numerical Operations,

R^2 = .30, F(6,125) = 9.10, p < .001, f^2 = .43, and Spelling, R^2 = .12, F(6,125) = 2.96, p = .010, f^2 = .14, were produced, explaining between 12% and 30% of the variance in academic achievement. Effect sizes ranged from large (for Numerical Operations) to medium (for Word Reading) to small (for Spelling). Adding the WMPT score at Step 2, produced a significant R^2 change in Word Reading, ΔR^2 = .03, F(1,124) = 5.00, p = .027, and Numerical Operations, ΔR^2 =.16, F(1,124) = 36.69, p < .001, but not in Spelling, ΔR^2 =.02, F(1,124) = 2.95, p =.089.

In the final models, the proportion of variance explained was highest for Numerical Operations, R^2 = .46, F(7,124) = 15.27, p < .001, f^2 = .85, followed by Word Reading, R^2 = .25, F(7,124) = 5.85, p < .001, f^2 = .33) and Spelling, R^2 = .15, F(7,124) = 3.00, p < .001, f^2 = .18, corresponding to large, medium and small effect sizes, respectively (Cohen, 1992). WMPT and AWMA Spatial Recall scores made significant contributions to the prediction of Numerical Operations. WMPT, AWMA Digit Span and AWMA Listening Recall scores all made significant contributions to the prediction of Word Reading. For Spelling, while the final model was significant, none of the working memory measures made a significant independent contribution to the model. Gender and age did not make a significant contribution in any of the models.

The results of the hierarchical regressions indicate three main findings: 1) The WMPT and the AWMA each predicted a significant proportion of the variance in Word Reading, Numerical Operations and Spelling scores when either test was entered at Step 1; 2) The inclusion of both forms of working memory assessment in the model (i.e., at Step 2) significantly increased the explained variance in Word Reading and Numerical Operations, over and above the contribution made by either the WMPT or the AWMA alone; 3)

Including both the WMPT and the AWMA did not significantly increase the proportion of explained variance in Spelling scores beyond that explained be either measure alone.

Discussion

The aim of the present study was to compare children's performance on the WMPT with a well-established test battery, the AWMA. In addition to assessing convergent validity between the two working memory assessments, we examined how scores from a single-test measure (the WMPT) compared with scores from a multiple-test battery (the AWMA) in predicting performance on tests of academic achievement in reading, numeracy and spelling.

Relationship Between WMPT and AWMA Test Scores

The results showed a strong correlation between children's performance on the WMPT and the AWMA total score, indicating good convergent validity. Significant correlations were also observed between WMPT accuracy and scores on each of the four subtests of the AWMA. The strongest correlation between the WMPT and the AWMA subtests was for Listening Recall, which is assumed to assess verbal working memory. This suggests that although the stimuli in the WMPT are nonverbal (line drawings of animals familiar to children), the test is also tapping verbal working memory. It is possible this relationship is due to verbal coding/rehearsal of the swaps involved in performing the WMPT (e.g., "Swap 1 with 2"). If this explanation is correct, one might have expected the correlation between the WMPT and Digit Span to also be high, as performance on the Digit Span task requires verbal information to be held for brief periods of time. Instead, a weak, but significant correlation was observed. An alternative explanation is that the correlation between the WMPT and Listening Recall is primarily due to the requirements of the central executive to perform the task. The significant correlation between WMPT scores and AWMA Spatial Recall, which is also assumed to involve the central executive, is consistent with this explanation. The significant correlation between WMPT and Dot Matrix scores would be assumed to reflect the visuospatial nature of the WMPT task.

Each of the above explanations is based on the multi-component model of working memory. A unitary model of working memory might account for the correlations between the WMPT and all four AWMA subtests by assuming performance of the WMPT involves keeping the verbal requirements of the task (e.g., Swap 1 and 2) in the focus of attention while mentally swapping the visual images of the animals, which also need to be held in the focus of attention. As the number of swaps increases, the amount of information in the focus of attention also increases. When capacity is exceeded, performance declines.

Working Memory as a Predictor of Achievement in Reading, Numeracy and Spelling

The second aim of the present study was to investigate whether a single test, the WMPT, can predict academic achievement in reading, numeracy and spelling as well as a multi-test battery such as the AWMA. We have shown performance on the WMPT is correlated with each of the AWMA subtests. We now turn to how well each of these working memory assessments contributes to the prediction of academic achievement.

The results of the hierarchical regression analyses showed working memory made a significant contribution to the prediction of achievement in numeracy, explaining 46% of the variance in the final model. WMPT scores made the largest contribution to the model, followed by AWMA Spatial Recall, which is designed to assess visuospatial working memory. The finding of a relationship between working memory and achievement in mathematics is consistent with previous research (e.g., Alloway et al., 2009; Chalmers & Freeman, 2017; Dumontheil & Klingberg, 2012; Lee, Ning, & Goh, 2014, McLean & Hitch, 1999). The relationship between performance on a visuospatial working memory task and mathematical ability has also been reported previously (e.g., Dumontheil & Klingberg, 2012; McLean & Hitch, 1999; Szucs et al., 2013). In addition, Dumontheil and Klingberg found that verbal working memory predicted mathematical ability. In contrast, verbal working memory did not make a unique contribution to the prediction of numeracy in the present study, even though the zero-order correlation between Listening Recall (which is assumed to assess verbal working memory) and Numerical Operations was significant. The difference is likely due to the WMPT accounting for variance that might otherwise have been explained by Listening Recall scores.

Working memory was also a significant predictor of achievement in Word Reading, explaining 24% of the variance. Three predictors, WMPT and AWMA Digit Span and Listening Recall, made unique contributions to the final model. For Spelling, while the overall regression model was significant, none of the individual predictors made a unique contribution to the model. The contribution of the verbal working memory tasks (Digit Span and Listening Recall) to the prediction of reading is consistent with previous research showing working memory as a predictor of literacy skills (e.g., Jarvis & Gathercole, 2003; Siegal & Ryan, 1989; Swanson et al., 2010). Depending on one's perspective regarding the structure of working memory, the contribution of WMPT scores to the prediction of reading could be due to children using verbal coding/rehearsal to perform the mental swaps required in the task (i.e., the multi-component view) or to a more fundamental relationship between working memory capacity and reading ability.

The results of the regression analyses showed the WMPT, a single-test measure of working memory, performed at least as well as the AWMA, which consists of four subtests designed to assess the components of the multi-component model of memory, in predicting academic achievement in reading, numeracy and spelling. On their own (i.e., when either the WMPT or the AWMA was entered at Step 1), each of these measures of working memory accounted for a significant proportion of the variance in Word Reading, Numerical Operations and Spelling. For both Word Reading and Numerical Operations, the addition of the second working memory measure to the model (i.e., adding AWMA subtest scores to the base model consisting of WMPT scores, and vice versa, adding the WMPT score to the base model consisting of AWMA subtest scores) led to a significant increase in the proportion of explained variance. This suggests both working memory measures contribute unique information to the prediction of achievement in reading and numeracy. For Spelling, the WMPT and the AWMA on their own (i.e., when entered at Step 1) produced significant models, but there was no significant improvement in prediction when both were included in the model. While the final model was significant overall, neither the AWMA nor the WMPT made an independent contribution to the prediction of achievement in Spelling.

A Comparison of the Components of Working Memory Assessed by the WMPT and AWMA

The finding that the WMPT and AWMA assessments each produce independent contributions to the prediction of achievement in reading and numeracy suggests that either they assess different aspects of working memory, or that one or both of them assesses cognitive abilities in addition to working memory. We now present a brief analysis of the two working memory assessments in terms of their updating requirements (Miyake et al., 2000; Morris & Jones, 1990) and factors such as inhibiting irrelevant information (Hasher, Zacks, & May, 1999) and binding of content and context (Oberaurer, 2005) that have been proposed to underlie capacity limitations in working memory. Differences between the WMPT and AWMA assessments in terms of the long-term memory requirements and response type are also considered.

The WMPT and the two AWMA working memory tasks (i.e., Listening Recall and Spatial Recall) involve elements of updating. In the WMPT, the information to be updated consists of associative bindings linking each of the three pictures with its location after the swap has been mentally applied. The resulting representation must be held in working memory until either the next swap is applied or the trial ends. In the AWMA working memory tasks, the information to be updated is a single item (a word, in Listening Recall; a dot position, in Spatial Recall), which is added to the items currently in working memory. In this case, the binding between content and context is temporally based. The updating process in the WMPT swap trials is similar to mental arithmetic involving several steps, where the results of a previous step become irrelevant once the next step has been processed, and need not be retained in memory. In contrast, the AWMA working memory tasks (as well as the short-term memory tasks) require information from each step to be retained until the trial ends and a response is made. These processes may be more similar to those required for reading and spelling.

The WMPT and AWMA working memory tasks differ in the extent to which they require the inhibition of irrelevant information (Hasher et al., 1999). While both forms of working memory assessment require inhibition of irrelevant information, they differ in the information that must be inhibited within the steps of a given trial. In the WMPT, the items in a trial remain the same, but the participant is required to mentally change the position of each item (and form new item-context bindings). The process is repeated, up to a maximum of 4 swaps. Performance will be improved if the bindings created during the previous, now irrelevant, swap (or swaps) is inhibited once the bindings resulting from the current swap are formed. In the AWMA working memory tasks, while some interference from the irrelevant sentences (or shape rotations) may occur, the irrelevant information has no relationship to the to-be-remembered word and presumably requires comparatively less effort to inhibit.

As noted above, the WMPT and AMWA working memory tasks both involve bindings between content and context (Oberaurer, 2005), although they differ in terms of the type of information to be bound (picture and spatial location in WMPT; word (or dot position) and temporal order in Listening Recall (or Spatial Recall). If, as has been previously proposed, temporal information is automatically encoded (Hasher & Zachs, 1979), then the binding requirements would likely be higher for the WMPT than the AWMA.

A notable difference between the WMPT and AWMA Listening Recall task is in the extent to which long-term memory is used in performance of the tasks. The WMPT uses novel, nonverbal stimuli and brief verbal instructions, whereas the Listening Recall task requires access to lexical representations of words and their meanings stored in long-term memory in order to perform the processing task. The AWMA Spatial Recall task, like the WMPT, is less reliant on long-term lexical representations. The finding that WMPT and Spatial Recall scores predicted performance on Numerical Operations suggests that numeracy may be related more to processing capacity than to memory for mathematical facts.

Another difference between the two assessment tools is in response type. The WMPT involves a recognition decision (choosing the correct sequence from a set of alternatives), whereas the AWMA requires recall of the sequence of words (or dots or digits) from memory. While this difference is unlikely to relate to the tests' prediction of academic achievement, future research could compare the visual recognition test format of the WMPT with an oral recall version of the WMPT to investigate this possibility.

Practical Implications

The results of the regression analyses indicated working memory ability is a significant predictor of academic achievement in reading, numeracy and spelling. Children with a deficit in working memory are likely to experience difficulties in achieving in one or more of these academic abilities. It is important, therefore, that we have easy-to-administer assessments available to allow the early detection of a working memory deficit. In practical terms, the present results indicate that a single-test measure of working memory, the WMPT, can be used to predict academic achievement at least as well as a multi-test battery, the AWMA. In contrast to the AWMA, which requires one-on-one administration with the examiner entering the child's responses on each trial, the WMPT can be completed online, with minimal supervision. The WMPT therefore has the potential to be used as a brief, easy-to-administer screening test for the early detection of working memory deficits. The WMPT uses non-verbal stimuli (i.e., line drawings of familiar animals) and the task requirements are written in simple English (e.g., "Swap 1 and 2"), making it suitable for children from different language backgrounds and abilities.

In addition, we have shown that while either the WMPT of the AWMA on its own can predict academic achievement, they each provide unique information in the prediction of reading and numeracy, but not spelling. The latter result suggests either the WMPT or the AWMA would be equally useful to examine whether poor performance in spelling may be due to a deficit in working memory. While the statistical results suggest more information about a child's abilities in reading and numeracy may be gained by administering both assessments, one might ask if this is necessary. In answering this question, it is important to remember that either instrument on its own predicted achievement in reading and numeracy (i.e., when entered into the regression models at Step 1) and would therefore be suitable for use in the assessment of working memory in children. Based on our analysis of the components of working memory assessed by each instrument, the WMPT could be useful in assessing working memory problems related to numerical ability (including inhibiting irrelevant information and updating), whereas the AWMA could be used in assessing problems in reading (which is heavily reliant on retaining temporal order).

Limitations and Future Directions

While the present results provide preliminary support for the WMPT as an easy-to-administer instrument to assess working memory problems in children, further research is needed to understand what additional information is being provided by the inclusion of both working memory assessments in the prediction of academic achievement. Future research should also aim to determine the cost-benefit of administering both forms of assessment in educational and clinical settings.

Participants in the present study were children aged 8 to 11 years. Further research needs to examine whether the present results generalise to older and younger children. The child's participation in the research was dependent on permission being granted by the child's parent or guardian, as well as their school principal. The range of scores in the working memory and achievement tests suggest self-selection into the study is unlikely to have impacted the findings, however future research could aim to rule out this possibility. Finally, future research could examine the extent to which the WMPT and AWMA predict achievement in other areas, including class tests and national curriculum assessments.

In conclusion, the present results suggest that a single-test measure of working memory can perform at least as well as a multiple-test assessment in predicting academic achievement across three fundamental abilities, reading, numeracy and spelling. Further research is needed to establish the generality of these results in terms of other age groups and for other measures of academic achievement.

References

Alloway, T. P. (2007). *Automated Working Memory Assessment*. London: Pearson Assessment.

Alloway, T.P., Gathercole, S.E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80*, 606-621.

Alloway, T.P., Gathercole, S.E., & Pickering, S.J. (2006). Verbal and visuo-spatial short-term and working memory in children: Are they separable? *Child Development, 77,* 1698–1716.

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences, 21*(1), 133-137.

Baddeley, A.D. (1992). Working memory. *Science, 255*(5044)*,* 556-559.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423.

Baddeley, A.D. (2007). *Working memory, thought and action*. Oxford: Oxford University Press.

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Bull, R., Espy, K.A., & Wiebe, S.A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*(3*), 205-228.*

Cameron, S., Glyde, H., & Dillon, H. (2014). Comparison of two working memory test paradigms: Correlation with academic performance in school-aged children.

International Journal of School and Cognitive Psychology, 1(3), 1-6.

Chalmers, K. A., & Freeman, E. E. (2017). Working Memory Power Test for Children. Journal of Psychoeducational Assessment, 1-7. Article first published online: September 15, 2017. https://doi.org/10.1177/0734282917731458

Cohen, J (1992). A power primer. *Psychological Bulletin*. *112 (1)*: 155–159.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.). *Models of Working Memory* (pp. 62-101). Cambridge: Cambridge University Press.

Cowan N. (2005). *Working memory capacity.* Hove, UK: Psychology Press.

Dumontheil, I. & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex*, 22*,* 1078-1085.

Freeman, E. E., Karayanidis, F., & Chalmers, K. A. (2017). Metacognitive monitoring of working memory performance and its relationship to academic achievement in Grade 4 children. *Learning and Individual Differences, 57,* 58-64.

Hasher, L., & Zachs, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). Cambridge, MA: MIT Press.

Jarvis, H. L., & Gathercole, S. E. (2003). Verbal and nonverbal working memory and achievements on national curriculum tests at 11 and 14 years of age. *Educational and Child Psychology, 20,* 123–140.

Lee, K., Ning, F., & Goh, H.C. (2014). Interaction between cognitive and non-cognitive factors: the influences of academic goal orientation and working memory on mathematical performance. *Educational Psychology: An International Journal of Experimental Educational Psychology, 31*(1), 73-91.

Mammarella, I. C., Lucangeli, D., Cornoldi, C. (2010). Spatial working memory and arithmetic deficits in children with nonverbal learning difficulties. *Journal of Learning Disabilities, 43*(5), 455-468.

McLean, J.F. & Hitch, G.J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240-260.

Metcalfe, A. W., Ashkenazi, S., Rosenberg-Lee, M., & Menon, V. (2013). Fractionating the neural correlates of individual working memory components underlying arithmetic problem solving skills in children. *Developmental Cognitive Neuroscience, 6*, 162-175.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81,* 111–121.

Niedo, J., Abbott, R. D., & Berninger, V. W. (2014). Predicting levels of reading and writing achievement in typically developing English-speaking 2nd and 5th graders. *Learning and Individual Differences, 32*, 54–68.

Oberaurer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition. *Journal of Experimental Psychology: General, 134*(3), 368-387.

Passolunghi, M. C., & Mammarella, I. C. (2010). Spatial and visual working memory ability in children with difficulties in arithmetic word problem solving. *European Journal of Cognitive Psychology, 22*(6), 944-963.

Siegel, L.S. & Ryan, E.B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*(4), 973-980.

St Clair-Thompson, H.L. & Gathercole, S.E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory, *Quarterly Journal of Experimental Psychology, 59*(4), 745–759.

Swanson, H.L., Kehler, P., & Jerman, O. (2010). Working Memory, Strategy Knowledge, and Strategy Instruction in Children With Reading Disabilities. *Journal of Learning Difficulties, 43*(1), 24-47.

Szucs, D., Devine, A., Soltesz, F., Nobes, A, & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex, 49,* 2674-2688.

Wechsler, D. (2007). *Wechsler individual achievement test – second edition, Australian Abbreviated.* London: Pearson Assessment.

Zheng, X., Swanson, H. L., & Marcoulides, G. A. (2011). Working memory components as predictors of children's mathematical word problem solving. Journal of Experimental Child Psychology, *110*(4), 481-498.

Notes

[1] An additional child was tested but not included in the analyses due to a Word Reading score that was more than 3 SD below the sample mean.

[2] The WMPT has been developed by ebilities. It is currently available for use in research studies. Please contact the publishers at ask@ebilities.com for further information.

[3] The availability of a confidence rating for each trial allows the calculation of metacognitive indices relating accuracy and confidence in performance (see Freeman, Karayanidis & Chalmers, 2017).

[4] Children were randomly assigned to one of the six possible orders of presentation.

Acknowledgements

We would like to thank the children, parents and schools whose participation made this research possible.