

## Comparison of Two Working Memory Test Paradigms: Correlation with Academic Performance in School-Aged Children

Sharon Cameron<sup>1,2</sup>, Helen Glyde<sup>1,2</sup> and Harvey Dillon<sup>1,2</sup>

<sup>1</sup>National Acoustic Laboratories, Australian Hearing Hub, 16 University Avenue, Macquarie University, NSW, 2109, Australia.

<sup>2</sup>HEARing Cooperative Research Centre, 550 Swanston Street, Audiology, Hearing and Speech Sciences, The University of Melbourne, VIC, 3010, Australia.

\*Corresponding Author: Sharon Cameron, PhD, Senior Research Scientist, National Acoustic Laboratories, Australian Hearing Hub, 16 University Avenue, Macquarie University NSW 2109, Australia, Tel: +61 2 9412 6851; Fax: +61 2 9412 6769; E-mail: [Sharon.Cameron@nal.gov.au](mailto:Sharon.Cameron@nal.gov.au)

Received date: May 13, 2014; Accepted date: July 22, 2014; Published date: July 29, 2014

Copyright: © 2014 Cameron, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

The purpose of this study was to examine the relationship between two different working memory task paradigms and academic achievement. Participants were 202 Australian primary-school children who were assessed on the Complex Auditory Span Evaluation (CASE) - a dual-task paradigm - and a reverse digit span paradigm, the number memory reversed test (NMR). Performance was correlated against the participants' National Assessment Program - Literacy and Numeracy (NAPLAN) results. Both the CASE and NMR were significant predictors of academic ability in literacy and numeracy. Whereas there was a significant correlation between the CASE and NMR, the relationship was weak ( $r=0.18$ ,  $p=0.012$ ). It was concluded that, although both types of test are related to academic achievement, NMR and dual-task paradigm tasks may be differentially sensitive to the working memory abilities required in different real-world situations. This result has implications for use of such tasks to predict academic performance.

**Keywords:** Working memory; Complex span; Dual-task paradigm; Academic achievement

### Abbreviations

CASE: Complex Auditory Span Evaluation; NAPLAN: National Assessment Program - Literacy and Numeracy; NMR: Number memory reverse; TAPS-3: Test of Auditory Skills - Third Edition

### Introduction

Working memory is the term used to refer to a system responsible for temporarily storing information while that material or other information is being mentally manipulated [1,2]. It functions as a mental workspace that can be flexibly used to support everyday cognitive activities such as arithmetic or listening [3-4]. For example, a deficit in working memory may impact a listener's ability to link noun phrases to their thematic roles or keep numbers in mind as well as interim and final results when performing numerical calculations.

Working memory capacity is predominantly regarded as tapping a domain-general attentional-resource limitation, although specific working memory tasks may show small degrees of modality specificity in their storage demands [5]. When performing working memory tasks individuals maintain memory items by switching their attention rapidly from processing to storage while performing the concurrent task [6]. Working memory capacity varies between individuals [2], and individual differences in working memory have been shown to be associated with academic performance [7-10]. Working memory deficits typically present in the classroom as difficulties in following instructions, tracking place when listening, detection of target items in spoken or written text, writing and longer-term remembering [11]. Further, disruptions in sentence comprehension due to working

memory deficits can result in language learning problems in children with specific language impairment [12].

Often children with working memory deficits are labelled as inattentive and unmotivated [11]. They are also judged as being distractible, frequently failing to monitor the quality of their work, and lacking in creativity in solving complex problems [13]. In explanation of these behaviours it was suggested that working memory overload will lead to the loss of crucial task information, and this forgetting will compromise the child's chances of completing a task successfully [13]. If unidentified, such deficits can lead to low learning outcomes [7] with the potential for low self-esteem and may result in children leaving school at an earlier age than desirable. As such, tests that assess working memory may be important tools to identify children at risk of academic underachievement in the classroom.

Commercially available tests of working memory typically involve live-voice presentation of strings of digits (or letters) to be repeated back to the examiner in reverse order. These working memory tests can form part of a lengthy battery of cognitive tests administered by an educational psychologist or speech language pathologist. An example is the digits backwards and letter-number sequencing subtests of the Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV) [14], or the Number Memory Reverse (NMR) subtest Test of Auditory Skills - Third Edition (TAPS-3) [15]. Research has shown that in children, but not in adults, a reverse digit memory test is more highly correlated with dual-task memory-span tests than it is with short-term memory tests that require no mental processing of the words heard [16].

An alternative to the more simplistic reverse digit span tests are dual-task span tests. In these dual-task paradigms a memory span test is combined with a concurrent processing or distracting task. An example is the Competing Language Processing Task (CLPT) [17]. In this test three-word sentences (e.g. Pencils eat candy) are presented

over headphones. The listener is required to respond to the truth-value of each sentence and then recall as many sentence-final words as possible. The sentences increase in set size (from 2 to 6). Sixty-eight children aged 6 to 12 years were assessed on the CLPT, two tests of short-term memory and the Peabody Picture Vocabulary Test-Revised (PPVT-R) [18], a test of receptive vocabulary. A significant correlation ( $r = .63$ ) was found between performance on the CLPT and the PPVT [17]. In reviewing this study, it was posited that additional reliable measures of functional working memory appropriate to school-age children and adolescents need to be developed, with future research focussing on establishing links between working memory and specific aspects of language learning and processing [19].

Eighty-three children aged six to eight years were assessed on a number of short-term memory tests, backwards digit recall, and dual-task working memory tests and scores were compared to results on the UK Key Stage 1 National Curriculum Assessments [20]. MANOVAs revealed significant effects of group (low vs. high achievers on the National Curriculum Assessments) for the working memory tests ( $p < 0.001$ ). Three working memory tasks - backwards digit recall, counting recall and listening recall (a dual-task paradigm akin to the CLPT) - yielded significant differences on univariate F-tests ( $p < .001$ ). The authors concluded that working memory tests characterise children who are falling below national levels for age on national curriculum assessments and may play a useful role in screening children at risk for educational underachievement.

This conclusion was supported by a later large-scale screening study which demonstrated that out of 308 children who were categorized as having a memory deficit, 67 and 70 per cent scored more than one standard deviation below the mean (i.e. at the 16th percentile or lower) on standardized tests of reading and mathematics respectively [21]. For a subgroup of these children aged six years and older ( $n = 167$ ) a fixed-order hierarchical regression procedure was used to investigate the relationship between memory (measured as a composite score composed of performance on 12 visual and auditory short term and working memory tests), general ability (measured by a test of vocabulary and a test of block design) and learning (measured by tests of reading and mathematics). It was found that both general ability and memory skills shared a substantial amount of variance with learning, and both these skills uniquely predicted outcomes in reading and mathematics [21]. It was concluded that children with low working memory typically have poor academic progress, inattentive behaviour and forgetting. However, as the memory measures were entered as a composite, the contribution of individual categories of memory tests was not reported.

The primary goal of the present study was to expand on previous research [21] in a large scale Australian study. The relationship between academic achievement and working memory performance was investigated using two different types of working memory tasks - a reverse digit span test and a dual-task paradigm. The NMR subtest of the TAPS-3 was utilized as the reverse digit span task as it can be administered and scored as a stand-alone test. An Australian-accented, computer-based, adaptive dual-task - the Complex Auditory Span Evaluation (CASE) [20] - with an automated scoring function was developed specifically for this study. Academic achievement was determined from participants' National Assessment Program - Literacy and Numeracy (NAPLAN) results - a battery of tests administered annually to Australian primary school children. In contrast to previous research [19], the present study would examine the relationship between academic achievement and both of the

working memory tests independently in order to elucidate the differential contribution of the abilities captured by these two types of working memory tasks to classroom performance. It was hypothesized that the dual-task working memory test would assess skills more similar to those required for functioning in the classroom (such as making judgments of about things heard, while simultaneously remembering them) than to the highly non-natural task of repeating back numbers in reverse order. As such, it was hypothesized that the CASE would correlate more highly with NAPLAN score - particularly with literacy ability - than the NMR subtest of the TAPS-3. A secondary goal of the study was the development of an adaptive dual-task working memory test. The adaptive nature of the CASE test design ensured that working memory thresholds were determined quickly and accurately whilst avoiding floor and ceiling effects. The development of the CASE is described in the methods section.

## Method

Approval for the study discussed in this paper was granted from the Australian Hearing Human Research Ethics Committee and the Catholic Schools Office, Diocese of Broken Bay.

## Participants

Participants were recruited from three Catholic primary schools. Participating schools had average Index of Community Socio-Educational Advantage values similar to the national average. Children in Years 3 and 5, who had completed the National Assessment Program - Literacy and Numeracy (NAPLAN) in May 2011 were invited to attend. Children diagnosed with unmedicated Attention Deficit Hyperactivity Disorder (ADHD) were excluded from the study. Data were collected from a total of 202 children aged between 8;4 (years;months) and 12;4. There were 98 children in Year 3 (males = 44). Mean age was 9;2 (range 8;4 to 9;10). There were 104 children in Year 5 (males = 34). Mean age was 11;2 (range 10;2 to 12;4).

## Procedure

Testing was carried out in a quiet room in the participating schools between 9 am and 3 pm. Testing took approximately 20 minutes per child.

## Materials

The participants were evaluated on the following materials. The presentation order of assessment tasks was counterbalanced between participants.

**Dual-task working memory test:** The Complex Auditory Span Evaluation (CASE) [22], took approximately 15 minutes to complete. The CASE was developed in Excel using Visual Basic for Applications. The test is designed to be administered on a personal computer over the computer speakers or under headphones. Client data is entered on the main screen and results are displayed on that screen. Over a number of trials, the child hears a series of short, pre-recorded sentences ranging from one sentence to a maximum of ten sentences. The sentences in each trial can be inherently true (e.g. Birds fly in the air) or false (e.g. Penguins wear red pyjamas). After a sentence in the series has been presented the child is asked to judge whether that sentence is true or false. Once all the sentences in a particular trial have been presented and veracity noted by the child, the child is asked by the test administrator to repeat either the first or the last word of

each sentence in the series, in any order. The position of the target word (that is, first or last) is randomly assigned by the software.

The child is told how many sentences he or she will hear (i.e. one to ten) in a particular series prior to each trial. The instructions given to the child appear as Appendix A. When the test administrator selects the Start Test button on the main screen a dialogue box appears which displays the pre-recorded sentences to be presented and shows whether the first or last word is to be repeated in any particular trial (Figure 1). Whether or not a sentence is correctly judged by the child to be true or false does not impact on scoring. The judgment step is included to add an additional element of mental manipulation to the working memory task.

The number of sentences per trial is adjusted adaptively. If more than 50 per cent of target (first or last) words is correctly identified the number of sentences presented in the next trial is increased by one. If less than 50 per cent of target words is identified the number of sentences in the next trial is decreased by one. If exactly 50 per cent of the target words is identified the number of sentences in the next trial remains constant.

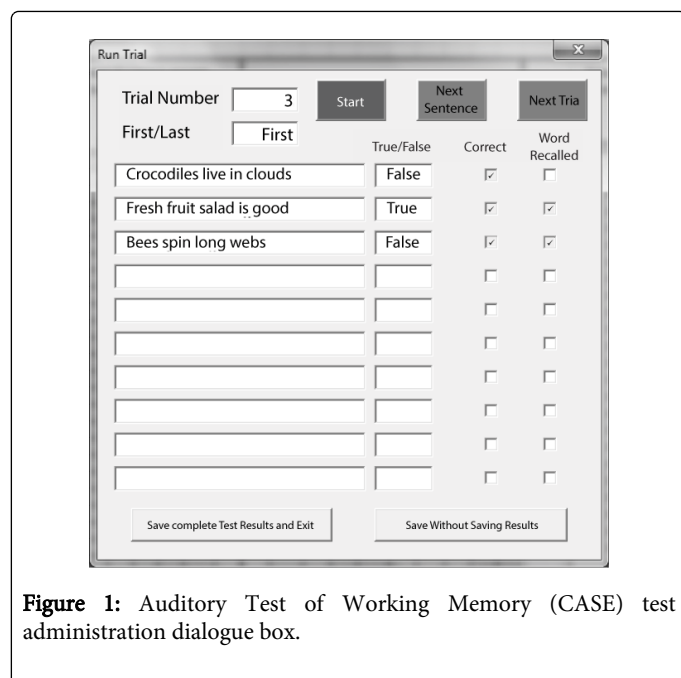


Figure 1: Auditory Test of Working Memory (CASE) test administration dialogue box.

The test commences with a period of practice which ends when the child reaches asymptotic performance, which is decided as follows. The average number of items recalled is calculated as the mean value of the scored trials, where the scored trials include all trials between the last trial administered and the trial that resulted in the smallest standard error of measurement (SEM) across all trials. Testing stops when at least 18 trials have been presented, or the SEM is less than 0.2 sentences and at least four trials have been scored. A maximum of 100 sentences appear in a fixed order. Results are recorded in a separate Excel spreadsheet. Percentage of true/false judgements correct and number of trials scored are included in the results spreadsheet.

The stimuli were 100, four and five word sentences developed by the research team. All semantic items were taken from The MacArthur-Bates Communicative Development Inventories [23] and are acquired by children aged 30 months of age (Figure 1). The sentences were spoken by a female speaker (the first author) and

recording took place in a chamber, anechoic above 50 Hz. The stimuli were recorded on a personal computer using Adobe Audition version 3.0, an M-AUDIO mobile pre USB audio interface and a Sennheiser ME64 cardioid microphone with a foam sock. The recordings were edited using Adobe Audition 3.0. Each sentence was saved as an individual speech file. These files were cut 5 ms before the start of the first word and 5 ms after the end of the last word. Each sentence file was then level normalized to have a root mean square (RMS) level of -22.0 dB re: digital full scale.

**Reverse digit span:** The numbers reversed subtest of the Test of Auditory Processing Skills – 3 (TAPS-3) [15] was utilized to assess how well the participant can retain and manipulate simple sequences of auditory information. The number sequences are read to the participant who is asked to repeat them in reversed order. The administrator is instructed to say each number clearly, in an even tone, with a pause between them. Sequences recalled in reversed sequence without errors receive a score of 2. Sequences that have the correct numbers, but are in the wrong order, receive a score of 1. If any number is omitted, or an incorrect number is substituted or inserted, a score of 0 is given. Testing is discontinued when the student makes three consecutive 0-point responses. Raw scores are converted to scaled scores (1 to 19, with a normative mean of 10 and a SD of 3) as a function of age. The procedure took approximately 5 milliseconds to complete.

**Literacy and Numeracy:** Academic performance was determined from participants' National Assessment Program – Literacy and Numeracy (NAPLAN) results. The NAPLAN is a battery of tests administered annually in May to Australian students in Years 3, 5, 7 and 9. NAPLAN results are available in September and the working memory testing for this study took place in November and December 2011. The literacy component of the NAPLAN assesses reading comprehension, writing (including text structure, vocabulary use, sentence structure, spelling and punctuation) and language conventions (spelling, grammar and punctuation). A combination of multiple choice answers and constructed responses are utilized. For the purposes of this study the participating schools, with the permission of the participants' primary caregiver, provided raw literacy and numeracy composite scores for each student. Performance is expressed as sample standard deviation units from the mean (z scores), calculated separately within each year group. NAPLAN results have been shown to have medium strength correlations with both receptive vocabulary, as measured by the Peabody Picture Vocabulary Test, and non-verbal ability, as measured by the Matrix Reasoning subtest of the WISC (r ranging from 0.3 to 0.55) [24].

## Results

Analyses were performed with Statistica 10.1. Table 1 documents the mean scores and SDs for each of the four measures. For the CASE, working memory capacity is reported as average number of items recalled. To enable results from Years 3 and 5 to be combined, scores for students in each school grade were separately standardized to a mean of zero and unity standard deviation. All measures are approximately normally distributed for the 202 participants in this study. There was, however, a significant departure from normality for the both the CASE ( $p=0.0001$ ) and TAPS-3 NMR ( $p=0.005$ ) based on the Shapiro-Wilk test of normality.

## Gender

The average z score on the CASE for males (0.17) was 0.25 population SDs higher than for females (-0.09), however a t-test revealed no significant difference between the groups ( $t[1,200] = -1.67, p=0.074$ ). Similarly, on the TAPS-3 NMR there was no significant difference between the average z score for males (0.15 SD) and females (-0.09 SD), ( $t[1, 200] = -1.68, p = 0.1$ ). There was also no significant difference in performance as a function of gender on the literacy component of the NAPLAN, with the average z score for males being -0.03 SD and 0.02 for females ( $t[1, 200] = 0.32, p = 0.746$ ). However, there was a significant difference between groups on the numeracy component ( $t[1, 200] = -2.52, p = 0.012$ ), with the average z scores for males being 0.22 SD compared to -0.14 for females.

Measure	Year 3		Year 5	
	Mean	SD	Mean	SD
CASE	4.2	1.3	5.2	1.6
TAPS-3 NMR	10.6	2.6	11.9	2.9
NAPLAN – Literacy	1858.2	260.9	2119.1	196.5
NAPLAN – Numeracy	1313.3	239.4	1547.3	177.2

Note. CASE = Complex Auditory Span Evaluation; NAPLAN = National Assessment Program - Literacy and Numeracy; TAPS-3 = Test of Auditory Processing Skills – Third Edition; SD = standard deviation

**Table 1:** Results on the CASE (number of sentences with 50 percent correct responses) and the NMR subtest of the TAPS-3 (scaled scores) calculated for the participants in Year 3 (n = 98) and Year 5 (n = 104). NAPLAN literacy and numeracy raw scores are also provided. Results are expressed as mean scores and SDs.

	CASE	TAPS-3 NMR	NAPLAN – Literacy	NAPLAN – Numeracy
CASE		0.177	0.225	0.248
TAPS-3 NMR	0.177		0.358	0.321
NAPLAN – Literacy	0.225	0.358		0.720
NAPLAN – Numeracy	0.248	0.321	0.720	

Note. CASE = Complex Auditory Span Evaluation; NAPLAN = National Assessment Program - Literacy and Numeracy; TAPS-3 = Test of Auditory Processing Skills – Third Edition; SD = standard deviation

**Table 2:** Correlations between results on the CASE, TAPS-3 NMR, and the two academic outcome measures. All correlations are significant to a  $p < 0.05$  level.

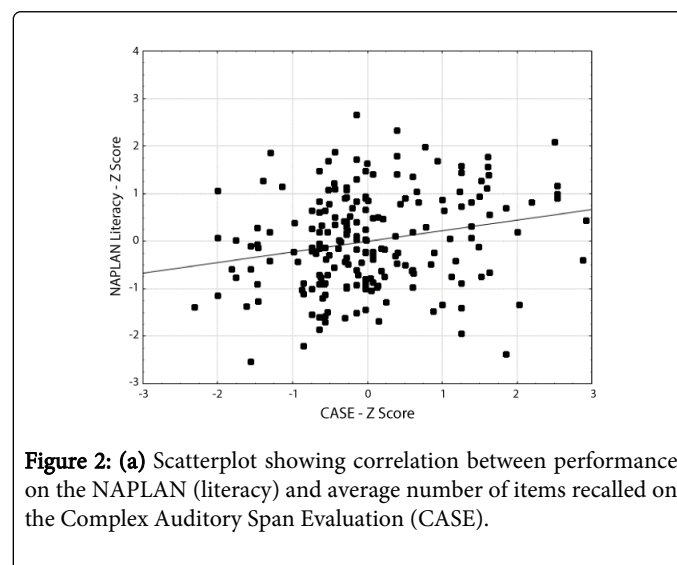
## Correlation analysis

Results from a product-moment correlation analysis are displayed in Table 2. All correlations were significant ( $p < 0.05$ ).

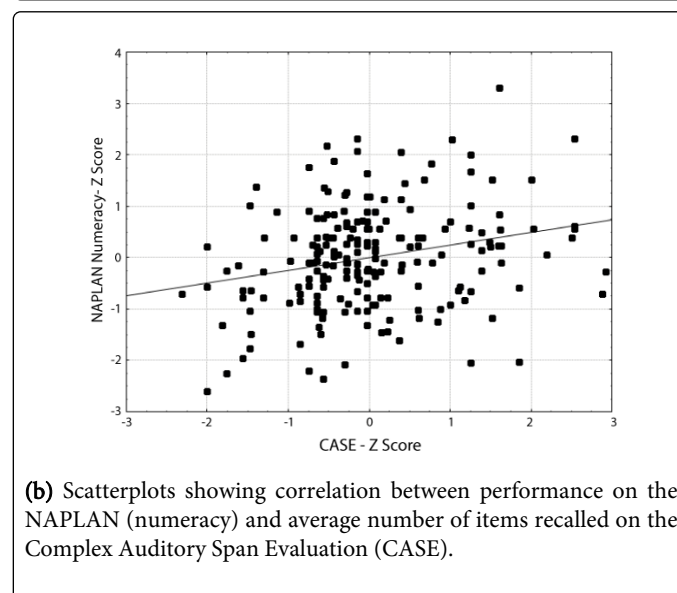
## Multiple regression analysis

Separate multiple regression analyses were conducted with each of NAPLAN literacy and NAPLAN numeracy z-scores as the dependent variables. For each analysis, predictor variables were gender, CASE z score and TAPS -3 NMR z score. CASE performance was a significant

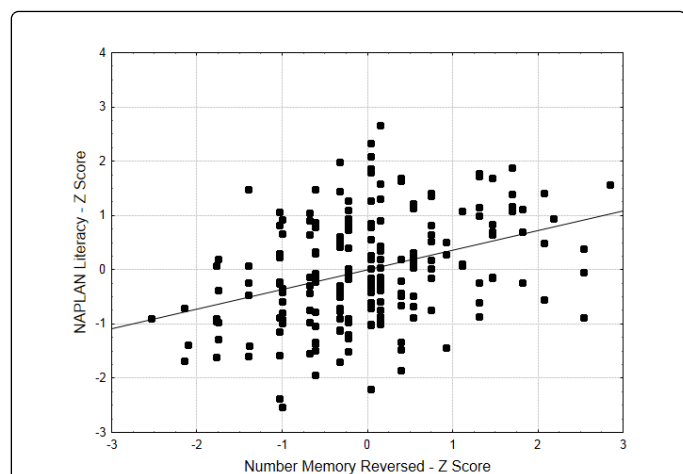
predictor of literacy achievement ( $t(198) = 2.64, p = 0.009, \text{Beta} = 0.176, \text{SE} = 0.067$ ). Performance on the CASE was also a significant predictor of numeracy achievement ( $t(198) = 2.67, p = 0.006, \text{Beta} = 0.184, \text{SE} = 0.067$ ). TAPS -3 NMR performance was a significant predictor of literacy achievement ( $t(198) = 5.07, p < 0.001, \text{Beta} = 0.337, \text{SE} = 0.066$ ). Performance on the NMR was also a significant predictor of numeracy achievement ( $t(198) = 4.10, p < 0.001, \text{Beta} = 0.274, \text{SE} = 0.067$ ). Given the standard errors of their respective beta values in the two regressions, the hypothesis that TAPS-3 NMR and CASE are equally good predictors of literacy or numeracy cannot be rejected. Scatterplots showing the relationship between NAPLAN literacy and numeracy results and performance on the CASE and TAPS-3 NMR are shown in Figure 2.



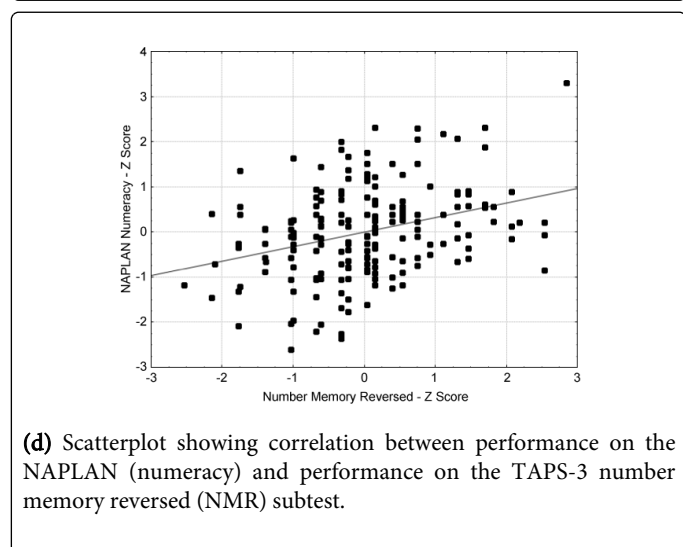
**Figure 2: (a)** Scatterplot showing correlation between performance on the NAPLAN (literacy) and average number of items recalled on the Complex Auditory Span Evaluation (CASE).



**(b)** Scatterplots showing correlation between performance on the NAPLAN (numeracy) and average number of items recalled on the Complex Auditory Span Evaluation (CASE).



(c) Scatterplot showing correlation between performance on the NAPLAN (literacy) and performance on the TAPS-3 number memory reversed (NMR) subtest.



(d) Scatterplot showing correlation between performance on the NAPLAN (numeracy) and performance on the TAPS-3 number memory reversed (NMR) subtest.

## Discussion

Performance on tests of working memory has been correlated with cognitive development as well as developmental cognitive disorders and as such, these tests commonly form part of psychometric and speech pathology evaluations. In Australia, working memory is most often evaluated using reverse digit span tests. The purpose of this study was to compare performance on a reverse digit span test and a verbal dual-task test of working memory and to investigate the relationship between these two types of working memory tests as well as their contribution to academic performance in Australian primary school-aged children. Reverse digit span was measured using the NMR subtest of the TAPS-3. A pre-recorded, adaptive dual-task sentence-based test, the CASE, was developed for this study to serve as the second working memory measure. Results on these tests were compared to participants' results on the National Assessment Program – Literacy and Numeracy (NAPLAN).

Both the CASE and the reverse digit span test were significantly correlated with academic ability in respect to literacy and numeracy as measured by the NAPLAN. TAPS-3 NMR appeared to be a better predictor of academic performance in this population than the CASE. Each 1 SD increase in NMR score resulted in a 0.34 SD increase in literacy score and a 0.27 SD increase in numeracy score. This compared to 0.18 SD increases in literacy and numeracy performance with CASE z score as the predictor variable. However, given the standard errors of their respective beta values in the two regressions (0.07 for all measures), we cannot reject the hypothesis that TAPS-3 NMR and CASE are equally good predictors of literacy or numeracy.

These results should be interpreted in light of the nature of the NAPLAN assessment which is a written test covering a wide range of academic skills (reading, writing, spelling, grammar and punctuation, and numeracy). As previously mentioned, the NAPLAN assessment has been shown to have medium strength correlations with both non-verbal cognitive ability and receptive vocabulary [24]. Therefore one cannot rule out the possibility that the relationships found here, between the two different measures of working memory and NAPLAN outcomes, are in some way mediated by these other cognitive abilities. Furthermore, although children may be capable of storing a particular amount of information in one situation, a demanding concurrent processing task will increase working memory demands and so may lead to memory failure [25]. An example would be listening to a story or a stream of information and then having to answer questions from any point in the discourse. In contrast, the majority of the NAPLAN test materials appear in written form allowing the student the opportunity to review information as often as required in order to respond to questions. It is possible that complex verbal dual-task paradigms such as the CASE may be a more sensitive than reverse digit span tasks to working memory deficits experienced during extended periods of classroom listening rather than written assessment tasks.

Colleagues investigating the incidence, mechanisms and perceptual consequences of auditory processing disorders in neurodegenerative diseases used the CASE in clinical trials with adults with Parkinson's disease (PD). Using scores on the Abbreviated Profile of Hearing Aid Benefit (APHAB) [26] as the dependent variable, the CASE was found to be a more sensitive predictor of self-reported degree of impairment ( $p = 0.011$ ) than a live voice reversed digit test ( $p = 0.044$ ). A multiple regression analysis, with auditory working memory, age, cognition and education levels as a predictor of APHAB score, found the CASE was still a stronger predictor of self-reported degree of impairment ( $p = 0.013$ ) than the reversed digit span test ( $p = 0.070$ ) (Rance; unpublished data).

Finally, a significant but weak correlation was found in the present study between performance on the CASE and TAPS-3 NMR ( $r = 0.18$ ). This result is somewhat surprising given previous research has suggested that both number memory reversed working memory tests and dual-task paradigms measure the same construct [2,16], which would typically result in a high correlation. However, the discrepancy between the findings reported here and previously published work could potentially be explained by protocol differences between the CASE and other dual-task memory paradigms. For instance, the CASE uses an adaptive testing approach whereas most other dual-task working memory tests increase the number of items systematically. Administration of the CASE also involves waiting until after all of the items in a trial have been presented before informing the participant whether they need to recall the first or last words, as opposed to tests

such as the CLPT where the participant knows they will be required to recall only the last item.

The weakness of the correlation between the TAPS-3 NMR and the CASE may suggest that the two tests are assessing different abilities, or that some other factor or factors not related to working memory is/are influencing performance on one or both of the tests, or that random measurement error on one or the other measures is limiting the correlation. Future research is planned whereby a group of children having difficulty following and comprehending information presented in the classroom will be assessed on the CASE and a live voice reverse digit span test with a test of auditory comprehension as the dependent variable.

## Conclusion

Whereas reverse digit span tests are more commonly utilized in Australia than the more complex dual-task paradigms in identifying children with academic delay due to working memory deficits, this study shows that both type of test can be used to predict academic achievement in school-age children. The weak correlation between the two tasks suggests that the reversed digit span and complex span tasks may be differentially sensitive to working memory abilities required in certain real-world situations. This has implications for the use of working memory tasks to assess academic performance and developmental cognitive disorders. For example, verbally presented, complex working memory tests such as the CASE - which involve listening, remember and making judgments - may be more sensitive and ecologically valid than reverse digit span tests in respect to assessing a child's ability to understand speech in the classroom over extended periods. As the majority of NAPLAN tests allow the child to review written information prior to making a judgment, further research is recommended in order to determine whether a particular type of working memory assessment task is more or less sensitive to individual differences in cognitive development and real-world skills in school-aged children using verbally-presented comprehension tasks as the dependent variable.

## Acknowledgements

The authors acknowledge the financial support of the HEARING CRC, established and supported under the Cooperative Research Centres Program – an initiative of the Australian Government, and the financial support of the Commonwealth Department of Health and Ageing. The authors would like to thank MsSujitaKanthan and Ms Anna Kania for their assistance in collection and management of data and assistance with development of the sentences for the CASE. We would also like to thank Mr John Seymour for his assistance in the recording of the speech stimuli for the CASE and Mr Mark Seeto for assistance with statistics.

## References

1. Baddeley A (2003) Working memory and language: an overview. *J CommunDisord* 36: 189-208.
2. Alloway TP (2006). How does working memory work in the classroom? *Educ Res Rev*, 1(4): 134-139.
3. Henry LA, MacLean M (2003). Relationships between working memory, expressive vocabulary and arithmetical reasoning in children with and without intellectual disabilities. *Educ Child Psychol*, 20(3): 51-63.
4. Schneider BA, Li L, Daneman M (2007) How competing speech interferes with speech comprehension in everyday listening situations. *J Am Acad Audiol* 18: 559-572.
5. Melby-Lervåg M, Hulme C (2013) Is working memory training effective? A meta-analytic review. *Dev Psychol* 49: 270-291.
6. Lépine R, Bernardin S, Barrouillet P (2005). Attention switching and working memory spans. *Eur J Cogn Psychol*, 17 (3): 329-345.
7. Alloway TP (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European Journal of Psychological Difficulties*, 25(2): 92-98.
8. Fry AF, Hale S (2000) Relationships among processing speed, working memory, and fluid intelligence in children. *Biol Psychol* 54: 1-34.
9. Gathercole SE, Alloway TP, Willis C, Adams AM (2006) Working memory in children with reading disabilities. *J Exp Child Psychol* 93: 265-281.
10. Towse J, Houston-Price C (2001) Combining representations in working memory: A brief report. *Brit J Dev Psychol*, 19: 319-324.
11. Gathercole S, Alloway T (2008) Working memory and classroom learning. In K. Thurman, K. Fiorello, (Eds.) *Cognitive Development in K-3 Classroom Learning: Research Applications*. Lawrence Erlbaum, NJ.
12. McGregor KK (2003) Working memory: an introduction to the conference proceedings. *J CommunDisord* 36: 185-188.
13. Gathercole S, Alloway T, Kirkwood H, Elliott J, Holmes J, et al. (2008) Attentional and executive function behaviours in children with poor working memory. *Learn Individ Differ*, 18: 214-223.
14. Wechsler D (2005). *Wechsler Intelligence Scale for Children*. (4th edn), The Psychological Corporation, San Antonio, TX.
15. Martin N, Brownell R (2005). *Test of Auditory Processing Skills*. (3rd edn), Academic Therapy Publications, Novato, CA.
16. St Clair-Thompson H (2010). Backwards digit recall: A measure of short-term memory or working memory. *Eur J Cogn Psychol*, 22(2): 286-296.
17. Gaulin CA, Campbell TF (1994) Procedure for assessing verbal working memory in normal school-age children: some preliminary data. *Percept Mot Skills* 79: 55-64.
18. Dunn LM, Dunn LM (1981). *Peabody Picture Vocabulary Test – Revised*. American Guidance Service, Circle Pines, MN.
19. Montgomery JW (2003) Working memory and comprehension in children with specific language impairment: what we know so far. *J CommunDisord* 36: 221-231.
20. Gathercole SE, Pickering SJ (2000) Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *Br J Educ Psychol* 70 : 177-194.
21. Alloway TP, Gathercole SE, Kirkwood H, Elliott J (2009) The cognitive and behavioral characteristics of children with low working memory. *Child Dev* 80: 606-621.
22. Cameron S, Dillon H (2011). *Complex Auditory Span Evaluation (Research Version 1.0.0)* [Computer software]. National Acoustic Laboratories, Sydney, NSW.
23. Fenson L, Dale P, Reznick JS, Thal D, Bates E, Hartung J, Pethick S, Reilly J (2007). *The MacArthur-Bates Communicative Development Inventories*. Paul H. Brookes Publishing Co, Baltimore, MA.
24. Daraganova, G, Edwards B, Siphthorp M (2013) *Growing up in Australia: The longitudinal study of Australian children (LSAC technical paper no. 8)*. Australian Institute of Family Studies, Melbourne.
25. Gathercole S, Alloway T (2007) *Understanding Working Memory. A Classroom Guide*. Harcourt Assessment, London.
26. Cox RM, Alexander GC (1995) The abbreviated profile of hearing aid benefit. *Ear Hear* 16: 176-186.