Evaluating the validity of the Automated Working Memory Assessment

Alloway, Tracy P., Gathercole, Susan, Kirkwood, Hannah, Elliott, Julian

Educational Psychology, 28 (7), December 2008, 725-734

http://dx.doi.org/10.1080/01443410802243828

The Working Memory Rating Scale:

A classroom-based behavioral assessment of working memory

Abstract

The aim of the present study was to investigate the potential of the Working Memory Rating Scale (WMRS), an observer-based rating scale that reflects behavioral difficulties of children with poor working memory. The findings indicate good internal reliability and adequate psychometric properties for use as a screening tool by teachers. Higher (i.e., more problematic) teacher ratings on the WMRS were associated with lower memory scores on direct assessments of working memory skills, as measured by the Automated Working Memory Assessment (AWMA) and the WISC-IV Working Memory Index. The use of the WMRS will allow educators to draw on their expertise in the classroom for early detection of children with working memory failures.

Keywords: working memory, behavior, Automated Working Memory Assessment, teacher ratings

Word count: 3061

The Working Memory Rating Scale: A classroom-based behavioral assessment of working memory

Working memory is the system that underlies the capacity to store and manipulate information for brief periods of time. According to leading models of working memory (Baddeley, 2000; Baddeley & Hitch, 1974; Baddeley & Logie, 1999), it can be distinguished from short-term memory as it involves both storage and processing of information, while short-term memory systems are specialised purely for the temporary storage of material within particular informational domains. Individual differences in working memory capacity have important consequences for children's ability to acquire knowledge and new skills (see Cowan & Alloway, in press). In the classroom, students frequently have to rely on working memory to perform a range of activities. Poor working memory leads to failures in simple tasks such as remembering classroom instructions (Engle, Cantor, & Carullo, 1991) to more complex activities involving storage and processing of information and keeping track of progress in difficult tasks (Gathercole & Alloway, 2008; Gathercole, Lamont, & Alloway, 2006).

Early identification of poor working memory skills in children is clearly desirable given the links between memory and learning. Existing teacher assessments of classroom behavior that include features of working memory are the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000) and the Conners' Teacher Rating Scale (CTRS; Conners, 2001). The BRIEF is a behavior checklist that yields scores relating to the following aspects of executive function: inhibition, shifting, emotional control, initiation, planning/ organization, organization of material, monitoring, and working memory. Given that the BRIEF measures various skills, one drawback is that it is a lengthy form and can overwhelm educators who have limited time to devote to early screening. While the CTRS is designed to identify attention failures on the basis of classroom behaviors, the Cognitive problems/Inattentive subscale includes one statement related to memory: ‘forgets things s/he has already learned’. However, this item confounds working memory with long-term memory.

According to some accounts, working memory is conceived as one of the executive functions responsible for goal-directed and problem-solving behavior (Pennington & Ozonoff, 1996). Although working memory shares a neuroanatomical association with the frontal lobes, current evidence suggests that in cognitive terms at least, it is distinct from other executive functions such as inhibition (Miyake et al., 2000) and has a separate link with learning (St Clair-Thompson & Gathercole, 2003). A more direct comparison of the dissociation between working memory and executive function, including sustained attention, is provided in a recent study by Holmes, Alloway, Gathercole, Hilton, Place, and Elliott (2008). Children with ADHD were compared with those with working memory deficits on a variety of cognitive and behavior measures. The findings indicate that the children with ADHD displayed elevated levels of impulsive and rule-breaking behavior in a range of executive function tasks and had higher rates of commission errors on a sustained attention task compared with the low working memory children.

On behavior measures, children with working memory deficits are typically judged by teachers to have poor attention span and high levels of distractibility but not the hyperactive/impulsive behavior characteristic of ADHD (Alloway, Gathercole, Kirkwood, & Elliott, in press-*a*; Gathercole et al., 2008; also Aronen, Vuontele, Steenari, Salmi, & Carlson, 2005). A recent comparison of performance of ADHD and low working memory children on the CTRS and the BRIEF indicated that these teacher ratings scales discriminated a significant proportion of the children with ADHD from those with low memory alone. While both groups exhibit behavioral problems in the classroom, they are characterized by differential attention profiles. The children with ADHD were rated as showing high rates of oppositional and hyperactive behaviors (CTRS), as well as problems with inhibitory control, shifting between activities, and emotional control (BRIEF). In contrast, the low working memory children were best characterized by problem behaviors related to working memory difficulties, including planning and organizing information (Alloway, Gathercole, Holmes, Place, Elliott, & Hilton, 2008).

Our own research programme in recent years has provided the opportunity for detailed and extensive observation of the classroom difficulties that characterise children who perform poorly on tests of working memory. A highly consistent profile of problems has emerged across observations of many different children, which include difficulties in following instructions, in remembering the detailed content of ongoing activities, in keeping track of progress through multi-step tasks, and in seeing through an activity to satisfactory completion (Gathercole & Alloway, 2008; Gathercole et al., 2006; see also, Engle, Carullo, & Collins, 1991). These observations provide an ideal starting point for the development of a behavior rating scale devised specifically to detect children who are at risk of having poor working memory and so of making the poor academic progress that typically accompanies this particular cognitive problem (Alloway et al., in press-*a*).

Although the existing behavior rating scales discussed above include some of these problems, they fail to capture the full behavioral profile typifying children with poor working memory.

The Working Memory Rating Scale (WMRS; Alloway, Gathercole, & Kirkwood, 2008) consists of 20 descriptions of these problem behaviors, and was designed to enable teachers to identify children at risk of the learning difficulties associated with working memory deficits. Key features of this tool are that it can be rapidly administered and simple to score. It focuses solely on working memory related problems in a single scale, and does not require any training in psychometric assessment prior to use. Furthermore, it can play a valuable role in familiarising teaching staff with classroom situations in which working memory failures frequently arise. Advice concerning ways of minimising the impact of working memory failures in the classroom is available both in the manual and in a variety of publications (Gathercole & Alloway, 2008; Gathercole, Elliott, & Alloway, 2008).

The present article provides new data on the internal consistency of the WMRS, and on its value in identifying children with very low scores on two direct assessments of working memory, the Automated Working Memory Assessment (AWMA, Alloway, 2007) and the Working Memory Index in the WISC-IV (Weschler, 2004). The AWMA was included as this is the only standardized tool for non-expert assessors such as teachers to screen their students for significant working memory problems. The automated mode of presentation and scoring of tasks provides consistency in presentation of stimuli across participants, thus reducing experimenter error. The WISC-IV Working Memory Index tests were also included as they provide an assessment of verbal working memory skills widely used by clinicians and psychologists. The inclusion of both the AWMA and the WISC-IV Working Memory Index allows for the direct comparison of behavior ratings with cognitive assessments of working memory skills.

Method

Participants

A total of 417 children from primary schools in the North-East of England participated in the study. Participating schools were selected on the basis of their performance on national assessments in reading, writing and mathematics, in order to achieve a representative sample achieving low, average, and high test results. The age of children ranged from 5.1 to 11.5 years (M=8.5 years, SD=20.5 months). Parental consent was obtained for each participating child.

A further 65 children with either low or average working memory skills were recruited in order to evaluate the extent to which the WMRS can discriminate between these groups on the basis of classroom behavior. The low working memory (LWM) group consisted of 27 children (20 boys; mean age=8.8 years, SD=20 months) identified via routine screening as having standard scores below 86 on both the listening recall and backwards digit recall tests from the Automated Working Memory Assessment (Alloway, 2007). A group of children with average verbal working memory scores (95-115) were selected from the same class as the LWM group (n=38; 16 boys; mean age=8.5 years, SD=25 months). Both these groups were screened as part of a larger study (Alloway et al., in press-*a*) and none were diagnosed with physical or sensory impairments.

Materials

The Working Memory Rating Scale (WMRS; Alloway et al., 2008) consists of 20 descriptions of behaviors that are characteristic of children with working memory deficits. Examples include: ‘The child raised his hand but when called upon, he had forgotten his response’; ‘She lost her place in a task with multiple steps’; and ‘The child had difficulty remaining on task’. Teachers are asked to rate how typical each behavior is of a target child, using a four-point scale ranging from (0) *not typical at all* to (1) *occasionally* to (2) *fairly typical* to (3) *very typical*. As, in the present study, the same classroom teacher rated both the low and average WM groups on the WMRS the potential for error due to the use of different raters was minimized

All 12 tests from the Automated Working Memory Assessment (AWMA, Alloway, 2007) were administered. In addition, two verbal working memory measures from the AWMA, listening recall and backward digit recall, were administered to the low and average WM groups. In the listening recall task, the child is presented with a series of spoken sentences, has to verify the sentence by stating ‘true’ or ‘false’ and recalls the final word for each sentence in sequence. In the backwards digit recall task, the child is required to recall a sequence of spoken digits in reverse order. Test reliability of the AWMA is reported in Alloway (2007) and test validity in Alloway, Gathercole, Kirkwood, and Elliott (in press-*b*).

The digit span (forward and backward) and letter-number sequencing tests from the WISC-IV (Weschler, 2004) were administered to the low and average WM groups. The raw scores were converted into scaled scores (*M*=10; *SD*=3). These were summed and converted into a standard score to represent the Working Memory Index (WMI). It is worth noting that the digit span score in the WISC-IV is a composite of forward and backward digit span. However, traditionally, forward digit recall is considered as a measure of verbal short-term memory as the processing load is minimal. In contrast, in backward digit recall the added requirement to recall the digits in reverse sequence imposes a substantial processing load (see Alloway, Gathercole, & Pickering, 2006, for further discussion).

Results

The data were screened for univariate outliers—scores more than 3.5 *SD* above or below the mean—on each WMRS item. Seven values out of the 9174 in the dataset met this criterion and were replaced with values corresponding to +/- 3.5 *SD* as appropriate. Cronbach's alpha across the whole sample was .978, establishing internal reliability of the scale.

<Table 1 here>

Table 1 provides descriptive statistics for the WMRS *T*-score (M=50, SD=10) and raw scores for the verbal and visuo-spatial working memory measures from the AWMA (n=417). A multivariate analyses of variance (MANOVA) was performed on the WMRS and the two AWMA scores, as a function of age in years (5 to 11 years) and gender. There was a significant effect of age, $F(4, 400)=34.15$, $p<.001$, but not of gender, $F<1$, and no significant interaction between age and gender, $F(4, 400)=1.72$, $p>.05$. The post-hoc analyses confirmed that there were no age effects for the WMRS *T*-score and the age effects were limited to the AWMA scores across all measures reflect the increasing memory capacity as children get older. Next, we investigated the constructs underlying the 20-item scale were explored using exploratory factor analysis with oblique rotation. A single factor accounted for 70.72% of the total variance. This suggests that the items are measuring a common factor thought to reflect working memory.

<Figure 1 here>

Confirmatory factor analysis was employed to evaluate the relative fit of a two-factor model tapping behavioral (measured by the WMRS) and cognitive aspects (measured by the AWMA) of working memory. A commonly used index of goodness of fit for such a model is the χ^2 statistic, which compares the degree to which the predicted covariances in the model differ from the observed covariances. A good fit is determined by small and nonsignificant χ^2 values. Because this statistic is sensitive to variances in sample sizes, with very large samples as in the present study even the best-fitting models frequently yield significant χ^2 values (Kline, 1998). Model adequacy was therefore evaluated using additional global fit indices that are more sensitive to model specification than to sample size (Jaccard & Wan, 1996; Kline, 1998). Fit indices such as the Comparative Fit Index (CFI; Bentler, 1990) and the Bollen Fit Index (IFI; Bollen, 1989) provide a further measure of fit computed by comparing the hypothesized model against a null model in which the relations between the latent variables are not specified and consequently are set at zero. Fit indices with values equal to or higher than .90 demonstrate a good fit.

In the measurement model tested, Factor 1 consisted of verbal and visuo-spatial working memory composite scores from the AWMA and Factor 2 consisted of the WMRS *T*-scores (M=50, SD=10) for the sum of all items. The path between the two factors was left free to co-vary (represented diagrammatically as bi-directional links) in the absence of justifiable assumptions concerning direction of causality (see Figure 1).

This model provided a good fit of the data, χ^2 = 2.09, *df* = 3, *p*>.10, and all fit indices were above .90 (CFI=.991, IFI=.991). The correlation between Factors 1 and 2 was .52, with 27% of their variance shared. This establishes a substantial relationship between the direct and rating-based assessments of working memory.

<Table 1 here>

Of interest was the relationship between the WMRS and the AWMA and the WISC-WMI in children with low and average working memory scores. Correlational analyses performed on children from both groups with age partialled out (*n*=65) established that high WMRS *T*-scores were significantly associated with low scores in the AWMA measures: listening recall (*r*=-.57) and backward digit recall (*r*=-.59); and the WISC-WMI (*r*=-.51; *p*<.001 in all cases).

In order to investigate the consistency of the subscale scores of the children within their groups, *T*-scores were banded into different categories. As there is no discrete point at which typical and atypical performance can be unequivocally distinguished, cumulative proportions over a range of values that represent different degrees of severity of low performance are presented. For the present purposes, values of one *SD* above the mean (*T*-scores >60) are viewed as indicative of a working memory deficit, with higher scores representing greater degrees of severity. The cumulative proportions of children obtaining *T*-scores below particular cut-off values were calculated (>60, >65, and >70).

The pattern of performance is very clear: there is a strong dissociation of teacher ratings that corresponds with working memory skills where the LWM children had high ratings and the AWM had low ratings. The majority of LWM children (67%) achieved elevated levels of problem behaviors (>60) on the WMRS. A smaller proportion also obtained markedly atypical scores: 49% and 33% on cut-off values of >65 and >70, respectively. This indicates that a high proportion of the LWM children in the present study exhibited the behaviors typically associated with working memory problems in the classroom. In contrast, only 11% of the AWM group achieved scores one *SD* above the mean (>60) on the WMRS, and none had scores in the atypical range (>65 and >70).

Discussion

The aim of the present study was to investigate the Working Memory Rating Scale (WMRS), an observer-based rating scale that assesses the classroom difficulties associated with poor working memory. The scale has good internal reliability and adequate psychometric properties for use as a diagnostic tool. Convergent validity was established with the two-factor measurement model that indicated the behavior ratings measure from the WMRS and the more conventional direct assessments of working memory provided by the AWMA measure are significantly related, establishing substantial construct validity. The convergence between teacher ratings on two direct standardized measures of working memory – the AWMA (Alloway, 2007) and the WISC-IV Working Memory Index (Weschler, 2006) indicate that it is a useful screening tool that will enable educators to draw on their classroom expertise to facilitate the early detection of children with working memory difficulties.

An important issue concerns the extent to which WMRS scores are predictive of other characteristics of working memory overload, such as high levels of inattention and mind-wandering (e.g., Alloway et al., in press-*a*; Gathercole et al., 2008; Kane et al., 2007). This was investigated in a recent study that compared teacher ratings on the

WMRS with the CTRS and BRIEF in low WM children (Alloway et al., 2008). There was a strong relationship between the WMRS score and the working memory subscale in the BRIEF and the Cognitive problems/Inattention subscale of the CTRS. When compared with a typically-developing group, the WMRS identified a greater proportion of children with working memory problems than both the BRIEF and CTRS. These data suggest that working memory, as measured by the WMRS, is linked to some aspects of attention; however, it is separable from other aspects linked with inhibition and self-regulation. It would be useful for future research to address the discriminant validity of the WMRS more directly.

Other issues such as temperament and motivation may also contribute to teacher ratings on the WMRS. For example, the negative halo effect has been reported when teachers rate children with ADHD as inattentive despite them not displaying these symptoms (Schachar, Sandberg, & Rutter, 1986; Stevens, Quittner, & Abikoff, 1998). It may be that children who are shy, bored, or unmotivated in the classroom may be incorrectly identified by the WMRS as having working memory problems. Further research is needed to investigate this possibility. However, it is worth noting that the purpose of the WMRS is as an initial screener that would lead to a clinical assessment to confirm working memory impairments. This follow-up assessment would be able to distinguish between children who are bored and unmotivated from those who have genuine working memory difficulties.

Currently, two major obstacles to the effective identification and management of working memory needs in the classroom are firstly, that working memory problems are difficult to detect from casual observation alone, and secondly, that there is an absence of suitable assessment tools that can be used by teachers to identify potential working memory problems. The WMRS enables teachers to conduct a more systematic means of screening for potential working memory problems than can be provided by unguided observation alone. This study represents a first step in establishing the reliability of this tool. Future research focusing on the validity of the WMRS, as well as teacher effects, would help to strengthen the efficacy of the WMRS in identifying potential working memory that merits detailed assessment.

References

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

Alloway, T.P., Gathercole, S.E., Holmes, J., Place, M., & Elliott, J. (2008). A comparison of sustained attention and behavioral problems in children with ADHD and children with poor working memory. *Manuscript submitted for publication.*

Alloway, T. P., Gathercole, S. E., & Kirkwood, H.J. (2008). *Working Memory Rating Scale.* London: Pearson Assessment.

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

Alloway, T.P., Gathercole, S.E, Kirkwood, H.J., & Elliott, J.E. (in press-*b*). Evaluating the validity of the Automated Working Memory Assessment. *Educational Psychology*.

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., Gathercole, S.E., Willis, C., & Adams, A.M. (2004). A structural analysis of working memory and related cognitive skills in early childhood. *Journal of Experimental Child Psychology, 87*, 85-106.

Aronen, E. T., Vuontela, V., Steenari, M.-R., Salmi, J., & Carlson, S. (2005). Working memory, psychiatric symptoms, and academic performance at school. *Neurobiology of Learning and Memory*, *83*, 33-42.

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

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47-90). New York: Academic Press.

Conners, K. (2001). *Conners' Teacher Rating Scale-Revised (S).* New York: Multi-Health Systems Inc.

Cowan, N., & Alloway, T.P. (in press). The development of working memory. In N. Cowan (Ed). *Development of Memory in Infancy and Childhood, 2nd edition*. Hove, England: Psychology Press.

Dollaghan, C., Campbell, T., Needleman, H. & Dunlosky, G. (1997). Reducing bias in language assessment: A processing-dependent measure. *Journal of Speech, Language and Hearing Research*, *40,* 519-525.

Engel, P.M.J., Heloisa Dos Santos, F., Gathercole, S.E. (in press). Are working memory measures free of socio-economic influence? *Journal of Speech, Language, and Hearing Research.*

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research*, *84*, 253-262.

Field, A. P. (2005). *Discovering statistics using SPSS (Second Edition).* London: Sage Press.

Gathercole, S.E, & Alloway, T.P. (2008). *Working memory and learning: A practical guide.* London: Sage Publications.

Gathercole, S. E., Alloway, T. P., Kirkwood, H. J., Elliott, J. G., Holmes, J., & Hilton, K. (2008). Attentional and executive function behaviors in children with poor working memory. *Learning and Individual Differences, 18*, 214-223.

Gathercole, S. E., Elliott, J. & Alloway, T. P. (2008). Identifying and supporting children with poor working memory. *Dyslexia Review, 19,* 4-8.

Gathercole, S. E., Lamont, E., & Alloway, T. P. (2006). Working memory in the classroom. In S. Pickering. (Ed.), *Working Memory and Education* (pp. 219-240). London: Elsevier Press.

Gioia, G.A., Isquith, P.K., Guy, S.C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function.* Florida, USA: Psychological Assessment Resources, Inc.

Holmes, J., Alloway, T. P., Gathercole, S. E., Hilton, K., Place, M., & Elliott, J. (2008). Working memory and executive function profiles of children with low working memory and children with ADHD. *Manuscript submitted for review.*

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

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., et al. (2005). Computerized training of working memory in children with ADHD - A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 177-186.

Kremen, W.S., Jacobsen, K.C., Xian, H., et al. (2007). Genetics of verbal working memory process: A twin study of middle-aged men. *Neuropsychology*, *21,* 569-580.

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.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 51-87.

Schachar, R., Sandberg, S., & Rutter, M. (1986). AGREEMENT BETWEEN TEACHERS RATINGS AND OBSERVATIONS OF HYPERACTIVITY, INATTENTIVENESS, AND DEFIANCE. *Journal of Abnormal Child Psychology, 14*, 331-345.

Stevens, J., Quittner, A. L., & Abikoff, H. (1998). Factors influencing elementary school teachers' ratings of ADHD and ODD behaviors. *Journal of Clinical Child Psychology, 27*, 406-414.

Wechsler, D. (2004). *Wechsler Scale of Intelligence (Fourth Edition)*. London: Pearson Assessment.

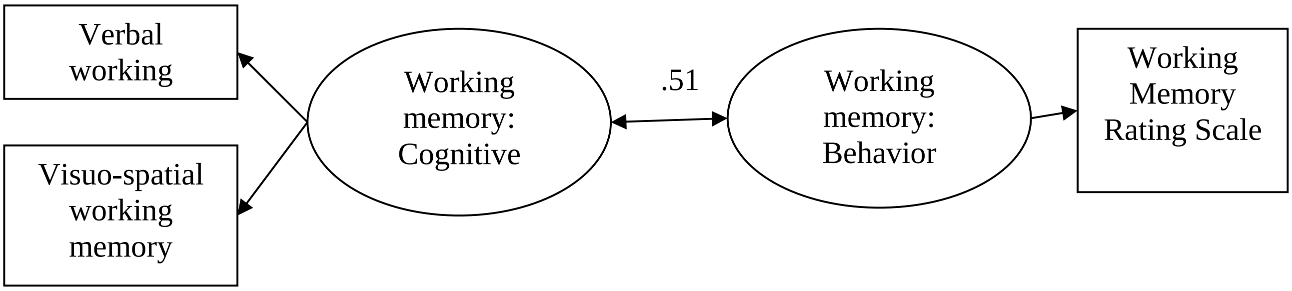
Table 1

Descriptive statistics for the WMRS and the AWMA for the whole sample (n=417) as a function of age and gender

		Male		Female	
Measures	Age (years)	Mean	SD	Mean	SD
WMRS*	5-6	50.81	9.53	49.36	10.06
	7	49.86	9.34	50.53	11.12
	8	46.16	7.64	53.03	10.67
	9	51.30	11.28	48.86	8.68
	10-11	50.69	10.46	49.17	9.47
AWMA: Verbal working memory	5-6	17.40	8.58	19.26	8.42
	7	28.68	11.10	26.50	9.43
	8	24.43	8.48	26.76	10.57
	9	35.09	13.97	35.78	10.36
	10-11	43.61	10.40	40.36	9.79
AWMA: visuo-spatial working memory	5-6	19.13	6.95	19.38	9.22
	7	30.80	13.74	28.70	9.46
	8	25.90	7.69	27.79	12.27
	9	36.81	12.18	37.69	12.21
	10-11	46.55	10.96	44.60	13.01

Note: * *T*-score: *M*=50 and *SD*=10.

Figure 1


Verbal
working
Visuo-spatial
working
memory
Working
memory:
Cognitive
.51
Working
memory:
Behavior
Working
Memory
Rating Scale