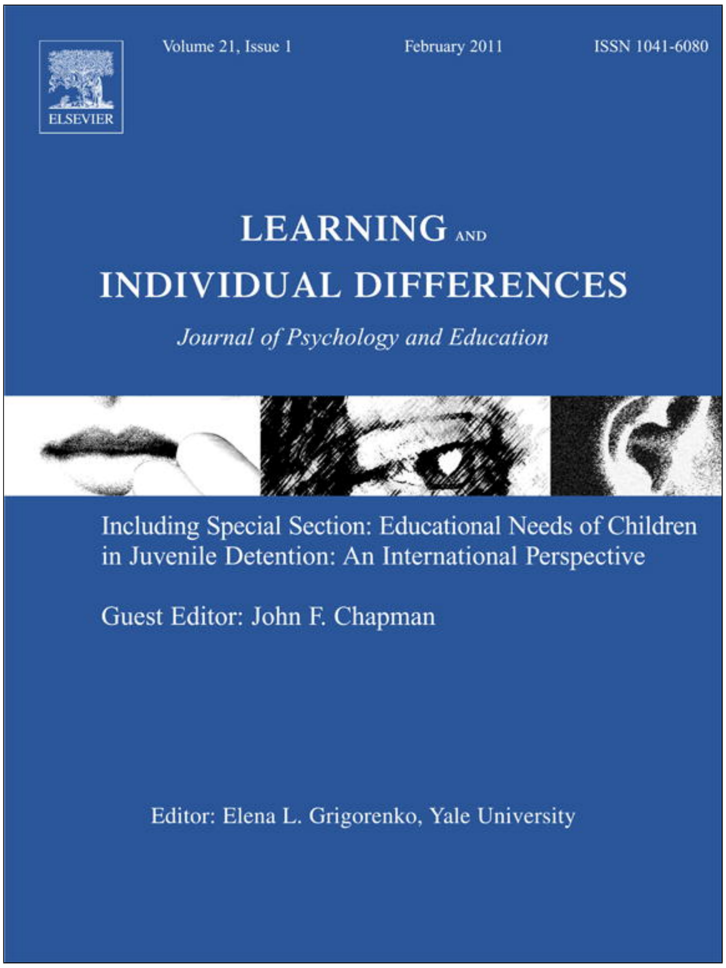
Volume 21, Issue 1
February 2011
ISSN 1041-6080
ELSEVIER
LEARNING AND
INDIVIDUAL DIFFERENCES
Journal of Psychology and Education
Including Special Section: Educational Needs of Children in Juvenile Detention: An International Perspective
Guest Editor: John F. Chapman
Editor: Elena L. Grigorenko, Yale University

Learning and Individual Differences 21 (2011) 78–84

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Assessing working memory in Spanish-speaking children: Automated Working Memory Assessment battery adaptation☆

Irene Injoque-Ricle [a,]*, Alejandra D. Calero [b], Tracy P. Alloway [c], Débora I. Burin [a]

[a] *Facultad de Psicología, Universidad de Buenos Aires–CONICET, Argentina*
[b] *Facultad de Psicología, Universidad de Buenos Aires, Argentina*
[c] *Psychology Department, University of Stirling, UK*

ARTICLE INFO

Article history:
Received 5 January 2010
Received in revised form 20 July 2010
Accepted 26 September 2010

Keywords:
Working memory
Children
AWMA
Psychometric

ABSTRACT

The Automated Working Memory Assessment battery was designed to assess verbal and visuospatial passive and active working memory processing in children and adolescents. The aim of this paper is to present the adaptation and validation of the AWMA battery to Argentinean Spanish-speaking children aged 6 to 11 years. Verbal subtests were adapted and pilot tested on a small sample ($n=26$). A validation study was conducted including 6-, 8- and 11-year-old children ($n=210$). All subtests presented an increase in difficulty as the number of to-be-remembered items raised, and showed high Cronbach's α values. Regarding validity, all subtests had medium to high and significant correlations among them, and with two external measures of working memory (Picture Span and Word Order) and an executive function task (Tower of London); correlations with Block Design were low and non-significant. We conclude that the adapted AWMA can be considered a valid and reliable battery of working memory in Argentinean Spanish-speaking children.

Evolving from early views about short-term memory, *working memory* refers to a set of temporary active contents and processes "involved in the control, regulation, and active maintenance of task-relevant information in the service of complex cognition" (Miyake & Shah, 1999, p. 450). The multi-component working memory model developed by Baddeley and colleagues (Baddeley, 1986, 1996; Baddeley & Hitch, 1974; Baddeley & Logie, 1999) distinguishes amodal central executive resources from two modality specific temporary retention systems, one verbally based (the phonological loop) and another that deals with visuospatial material (the visuospatial sketch-pad). This view is backed by convergent evidence from dual-task studies, neuropsychological cases, neuroimaging studies, developmental patterns, and psychometric approaches (Miyake & Shah, 1999). Both verbal and visuospatial components have limited duration and capacity, and seem to be composed of a passive buffer whose contents are refreshed by modality specific active maintenance mechanisms (Baddeley & Logie, 1999). The central executive component is characterized as a collection of attentional control, planning and monitoring, and retrieval resources (Baddeley, 1996).

Traditional measures of short-term memory span require subjects to immediately reproduce, in serial order, a sequence of simple stimuli such as digits, words, or visuospatial stimuli. In child assessment, examples of these type of tests can be found in the Digit Span subtest of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV, Wechsler, 2003) and in the subtests Word Order, Number Recall and Hand Movements of the Kaufman Assessment Battery for Children-Second Edition (KABC-II, Kaufman & Kaufman, 2004). These tests seem to tap the maintenance components of working memory, and they generally exhibit low correlations with complex cognitive performance (Baddeley & Logie, 1999; Dixon, LeFevre & Twilley, 1988; Fürst & Hitch, 2000; Gathercole & Baddeley, 1993) or academic achievement (Alloway et al., 2005; Gathercole & Pickering, 2000). Nevertheless, they have clinical relevance since they can detect specific phonological loop or visuospatial deficits; for example, short-term language retention has been related to specific neuropsychological language impairments (e.g. Vallar & Shallice, 1990); and to the acquisition of vocabulary and beginning writing in small children (Baddeley, Gathercole & Papagno, 1998). On the other hand, complex working memory span tasks require simultaneous effortful processing such as verifying grammar in sentences or solving arithmetic problems, while at the same time keeping in memory some information for subsequent recall (Conway et al., 2005). The number of memory elements is varied systematically, so that capacity or span can be computed for each subject. An example in children's batteries is Letter–Number Sequencing subtest of the WISC-IV (Wechsler, 2003). Another frequent paradigm, termed Sentence Span or Reading Span, asks for sentence verification (true or false) and subsequent

☆ The Automated Working Memory Assessment battery was translated and reproduced by Permission. Copyright © 2007 by Harcourt Assessment. Spanish Translation Copyright © 2007 by Harcourt Assessment. All rights reserved.

* Corresponding author. Instituto de Investigaciones, Facultad de Psicología, Universidad de Buenos Aires. Av. Independencia 3056, 3° piso (C1425AAM), Ciudad Autónoma de Buenos Aires, Argentina. Tel.: +54 11 4957 5886.
E-mail address: iinjoque@psi.uba.ar (I. Injoque-Ricle).

1041-6080/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.lindif.2010.09.012

recall of the last words of each sentence. The number of sentences is increased in each trial. The working memory span would be equal to the maximum number of sentences' last words the subject can recall correctly. Different approaches have emphasized diverse sources of capacity limitations, such as the amount of activation available for temporary retention and computations (Just & Carpenter, 1992), general processing speed (Hale & Fry, 2000), a general attention capacity coupled with strategic activation of long-term-memory relevant knowledge (Engle, Kane & Tuholski, 1999; Kane, Conway, Hambrick & Engle, 2007), specialized resources — for example, for syntactic and sentence comprehension (Caplan & Waters, 1999), or previous knowledge and experience with certain tasks and materials (Long & Prat, 2008; MacDonald & Christiansen, 2002). Although there are diverse interpretations of the source of individual differences at the "micro-analytic level" (Kane et al., 2007), working memory measures are taken as an index of a central resource, an overall working memory capacity, predictive of academic achievement and skills. In contrast with simple span measures, working memory span ones generally have shown high correlations with complex cognition such as language comprehension, math performance, abstract reasoning or "g", and acquisition of various abilities (e.g. Ackerman, Beier & Boyle, 2005; Conway et al., 2005; Kane et al., 2007). In children, working memory has been related to academic achievement (Alloway & Alloway, 2009; Bull & Scerif, 2001; De Jong, 1998; Gathercole, Brown & Pickering, 2003; Hale & Fry, 2000; Pickering & Gathercole, 2004). Overall, then, in educational and clinical settings it is useful to have measures for short-term and complex working memory spans.

The Automated Working Memory Assessment (AWMA, Alloway, 2007a) was developed to specifically address working memory performance according to the multi-component model (Baddeley, 1986, 1996; Baddeley & Logie, 1999). It consists of twelve computer administrated tests, half verbal, half visuospatial in materials and processing. They all require sequential recall of items, half of them also demanding concurrent processing upon presentation of the to-be-remembered material. In addition to its theoretical foundation, the AWMA has shown sound psychometric properties in large-scale samples (Alloway, Gathercole & Pickering, 2006), and convergent validity with concurrent clinical measures of working memory deficits (Alloway, Gathercole, Kirkwood & Elliott, 2008). It has clinical relevance since it demonstrated particular patterns of Working Memory impairment in children with Developmental Coordination Disorder (Alloway, 2007b; Alloway & Archibald, 2008; Alloway, Rajendran, & Archibald, 2009), Specific Language Impairment (Alloway & Archibald, 2008; Alloway, Rajendran et al., 2009), Attention-Deficit/Hyperactivity Disorder (Alloway, Rajendran et al., 2009), Asperger syndrome (Alloway, Rajendran et al., 2009), also contributed to delineate a specific Working Memory Impairment (Alloway, Gathercole, Kirkwood & Elliot, 2009; Gathercole et al., 2008). It has also shown predictive success for academic achievement (Alloway et al., 2005; Alloway & Alloway, 2009; Gathercole et al., 2003).

This paper presents an adaptation and validation of the AWMA battery to Argentinean Spanish-speaking children aged 6 to 11 years. Given that the AWMA is an English battery, all the instructions, and the content of three verbal tests had to be translated and adapted to Spanish. The adaptation of verbal tests sought psychometric (not literal) equivalence, which required taking into account aspects of phonology, orthography, syntax, semantics, and communicational context (for example, word frequency). In addition, since Spanish has marked regional variations, to select the verbal stimuli we employed a normative Spanish dictionary (LEXESP) constructed upon a written corpus of different literary genres, expository (popular science, editorial essays) texts, and press articles (news, sports) from diverse Spanish-speaking countries, not only Spain or a particular Hispano-American country (Sebastián Galles, Martí Antonín, Carreiras Valiña & Cuetos Vega, 2000). So then, although this study validated the battery for Spanish speakers in Argentina, it is aimed as the first step in constructing a neutral battery for Spanish speakers.

1. Adaptation

In the first place, the battery was translated to Spanish by a bilingual psychologist. Two psychologists and a speech pathologist performed the successive adaptations, which concerned the instructions, but mainly the three verbal tests: Word Recall, Nonword Recall, and Listening Recall.

For the Word Recall test, after an initial literal translation from English to Spanish, words that had more than two syllables were replaced (avoiding the word length effect) by monosyllabic, or 3–4 letter disyllabic, medium- or high-frequency words. Medium-frequency sets comprised words with a frequency between 16 and 99 words per million and high frequency sets comprised words with a frequency greater than 100 words per million according to the norms of Sebastián Galles et al. (2000). The disyllabic words had to follow the most frequent Spanish phonological structure, consonant–vowel–consonant–vowel. In English there may be only one word for adjectives and common nouns, but in Spanish the word varies according to its gender and number. Regarding verbs, in Spanish they are conjugated for the various pronouns. Given such grammatical considerations, only singular, common nouns were included. For trials composed by more than one word, care was taken to exclude phonological and semantic similarities between them, to avoid facilitation or disturbance of recall (phonological similarity effect, semantic associations). Furthermore, trials that had more than two items were formed by one monosyllabic and one or more disyllabic words. Table 1 summarizes these criteria.

Once the Word Recall test was designed, items for the Nonword Recall test were derived. Letters of the Word Recall items were changed, rendering the nonwords. Care was taken to keep the formal (phonological and morphological) structure of items of the former in the latter. A Spanish Dictionary (RAE, 2001) was consulted to ensure that they were nonwords. In addition, nonwords that formed frequent (or culturally known) words in other languages such as English, Portuguese, Italian or French, were modified. Again, in multi-item trials care was taken to include only one monosyllabic nonword. Table 1 summarizes these criteria.

For the Listening Recall test, after the initial translation, sentences were modified in various ways. Some sentences resulted more syntactically complex after the translation (for example, including nominal constructions instead of a single word, or requiring multiple congruencies of gender and number) and were corrected to comply with a simpler syntactical structure. Total length of the sentence was also checked. As for the to-be-remembered words, they were adjusted

Table 1
Word Recall, Nonword Recall, and Listening Recall adaptation criteria.

Word Recall	1. Only mono- or disyllabic common nouns included
	2. Word frequency medium to high
	3. Phonological structure of the disyllabic words: C–V–C–V
	4. Phonological and semantic similarities avoided in multi-item trials
	5. Only one monosyllabic item in multi-item trials
Nonword Recall	1. Items do not appear in Spanish RAE dictionary
	2. Items forming culturally known or frequent words in English, French, Portuguese or Italian not included
	3. Non words follow the same phonological and morphological rules as the Word Recall subtest
	4. Multi-item trials follow the same rules as the Word Recall subtest
Listening Recall	1. Only mono- or disyllabic words included
	2. Word frequency medium to high
	3. Sentences with simple syntactic structure: S–V–O active voice.

to be mono- or disyllabic. Also, care was taken to include only nouns (excluding conjugated verbs and adjectives). Table 1 summarizes these criteria.

1.2. Experiment 1

A pilot study was conducted, to evaluate comprehension of the instructions and the adapted verbal tests' functioning.

1.2.1. Method

1.2.1.1. Participants. Twenty six 13-year-old (21 girls and 5 boys) monolingual Spanish-speaking children, belonging to a middle-class public school in Buenos Aires city participated in the study. We selected children of this age because we needed a sample old enough to complete as much as new verbal items as possible. Children with a diagnosed psychiatric or neurological condition, language or hearing impairment, or a history of academic failure (repeating course) were excluded from the study. This information was provided by school reports. The adapted verbal tests were administered to all the twenty six children, with a mean age of 162.39 months ($SD = 4.35$). The complete battery to evaluate the comprehension of the instructions was administered to twenty of the twenty-six, with a mean age of 162.70 months ($SD = 4.24$). Their parents gave written informed consent.

1.2.2. Materials and procedure

All twelve AWMA tests were administered. *Digit Recall*, *Word Recall* and *NonWord Recall*, where the children have to recall a sequence of numbers, words or nonwords, respectively, in the correct order, that were presented verbally through a computer. *Dot Matrix*, where a sequence of red dots appears on a matrix and the child has to point the squares of the matrix on which each dot appeared, in the same order. *Block Recall*, which has the same procedure as the Dot Matrix, but the children see on the computer screen a board with nine cubes, located randomly, and some of these cubes are touched by an examiner. *Mazes Memory*, where the children see a two-dimensional maze with a path drawn on it, and have to recall the path, tracing it with their finger. *Listening Recall*, in which the children listen to a series of short sentences and have to decide whether they are true or false, and then have to recall the last word of each sentence on the exact order. *Counting Recall*, a test in which a series of dots and arrows appears on the computer screen and the children have to count aloud one by one all the dots, and finally, recall the total number of dots that appear on each trial, on the correct order. *Backward Digit Recall*, where sequences of numbers are verbally presented and the children have to recall them in the reverse order. *Odd One Out*, in which sets of three shapes in a three square matrix are shown on the computer screen, two are the same and the third one is different. The children have to indicate which one is the odd one out, and after a certain amount of sets is presented, have to indicate the place in which the odd figure was, in the exact order. *Mister X*, where sets of two figures of a men are presented, one with a yellow hat and the other with a blue one, and both with a red dot in one of their hands. Also, the Mr. X with the blue hat can appear rotated in six possible positions. The child has to say whether they have the dot in the same hand or in a different one, and then, in a picture with six compass points, has to point the location of the red dot in sequence. Finally, *Spatial Span*, a test in which sets of two arbitrary shapes are presented, one with a red dot on top of it, and one of them can be rotated in three possible positions. First, the children have to say if the shape with the red dot is the same or the opposite than the one without the dot, and then, they have to point in sequence the location of the dot in a three compass points.

Children completed the AWMA in an individual session lasting approximately 40 min, in an office at their school. AWMA was administered without "move-on" and "discontinuation" criteria, to be able to analyze item functioning.

1.2.3. Results and discussion

Instructions were well understood by all participants. Descriptive statistics for the AWMA tests are shown in Table 2.

Difficulty level and internal consistency of the three adapted verbal tests was analyzed. In all three tests, difficulty increased as a function of trial size, that is, the number of to-be-remembered items, as can be seen in Fig. 1. Internal consistency of these three tests resulted acceptable (Word Recall: Cronbach's $\alpha = .75$; Nonword Recall: Cronbach's $\alpha = .71$; Listening Recall: Cronbach's $\alpha = .75$).

The pilot study showed that children had a good understanding of the instructions, and that the three adapted verbal tests had acceptable psychometric properties, although the small sample size limits this conclusion. Test characteristics were further examined in the validation study.

1.3. Experiment 2

The adapted AWMA was administered to a larger sample of children, along with other measures of cognitive functions, to obtain evidence about the battery's psychometric characteristics (item difficulty, internal consistency). Convergent validity was examined with two working memory tests, Picture Span (PS; Injoque-Ricle & Burin, 2008a), and Word Order (WO; Kaufman, 1983), and an executive function test, Tower of London (TOL; Injoque-Ricle & Burin, 2008b; Shallice, 1982). Regarding discriminant validity, working memory has been shown to be positively associated with a large number of cognitive processes and abilities but seldom with visuo-constructive ability (Conway et al., 2005; Hale & Fry, 2000; Kane et al., 2007; Luciana, Conklin, Hooper & Yarger, 2005). In addition, in the WISC-IV validation studies, Digit Span and Letter–Number Sequencing exhibited their lowest correlation with a

Table 2
Descriptive statistics for AWMA adaptation pilot study.

		Female	Male							
	n	*n*	*n*	Mean	Median	Standard deviation	Minimum	Maximum	Skewness	Kurtosis
Digit Recall	20	15	5	27.30	28.50	3.33	22	32	−.372	−1.167
Word Recall	26	21	5	22.35	21.50	3.32	18	31	.891	.742
Nonword Recall	26	21	5	15.12	14.50	3.01	9	23	.518	.672
Dot Matrix	20	15	5	29.60	30.00	3.92	23	35	−.371	−.828
Block Recall	20	15	5	29.50	30.50	2.27	26	32	−.426	−1.669
Mazes Memory	20	15	5	26.60	26.50	3.47	19	32	−.815	2.192
Listening Recall	26	21	5	14.40	13.50	3.07	10	21	.576	−.359
Counting Recall	20	15	5	18.90	19.00	3.70	12	25	−.298	.436
Backward Digit Recall	20	15	5	15.30	16.00	2.54	11	20	−.064	.692
Odd One Out	20	15	5	21.40	22.50	4.22	13	27	−.737	.153
Mr. X	20	15	5	14.00	14.50	5.03	7	24	.471	.472
Spatial Span	20	15	5	19.40	19.00	5.56	11	32	1.017	2.761

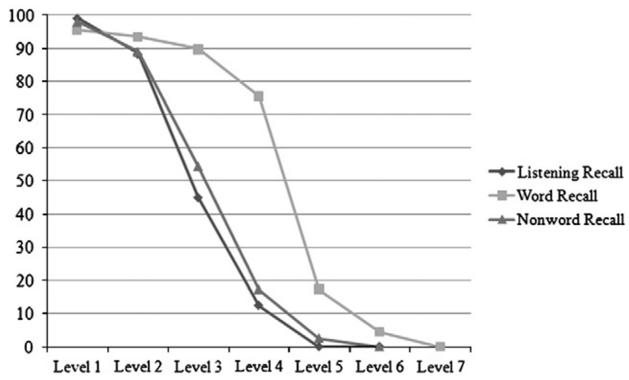

Fig. 1. Percent correct responses per trial size in Listening Recall, Word Recall and Nonword Recall tests.

Table 3
Age and gender distribution.

	Age					
	6		8		11	
	n	%	*n*	%	*N*	%
Female	33	47.1	39	55.7	50	71.4
Male	37	52.9	31	44.3	20	28.6
Total	70	100	70	100	70	100

test for this factor, Block Design, and with a perceptual one (Wechsler, 2003). Therefore, to compute discriminant validity, Block Design from WISC-III (Wechsler, 1994), which is the last validated WISC local version, was included.

1.3.1. Method

1.3.1.1. Participants. Two-hundred and ten monolingual Spanish-speaking children participated in the study. They belonged to three age groups ($n = 70$ ss. each): 6-, 8-, and 11-years old. The younger group had a mean age of 78.90 months ($SD = 3.30$). The 8-year old group had a mean age of 101.17 months ($SD = 3.35$). The older group had a mean age of 137.95 months ($SD = 2.79$). Distribution of gender by group is shown in Table 3.

Children attended, and were tested at, two middle-class public elementary-level schools in the city of Buenos Aires. Parents were invited to an informative meeting about the study, after which informed consent was solicited in order to allow their children to participate. A written devolution was provided for each child's parents.

Possible participants with a diagnosed psychiatric or neurological condition, language or hearing impairment, or a history of academic failure (repeating course) were excluded from the study. This information was provided by the school registers.

1.3.2. Materials and procedure

Children completed the previously adapted AWMA, two working memory tests, Picture Span (PS; Injoque-Ricle & Burin, 2008a), and Word Order (WO; Kaufman, 1983), an executive functions test, Tower of London (TOL; Shallice, 1982; Injoque-Ricle & Burin, 2008b), and a spatial processing test, Block Design from WISC-III (BD, Wechsler, 1994).

Each child was tested in a quiet room inside the school he or she attended, in 2 sessions lasting approximately 40 min each.

1.3.3. Results

Table 4 and Fig. 2 show the AWMA tests' distributions. Concurrent cognitive measures' distributions are also shown in Table 6. The box-plots, and the relative low values of kurtosis and skewness, suggest a normal distribution for AWMA tests, except that Digit Recall, Backward Digit Recall, and Block Recall seem to be positively skewed. The graphs point to several possible outliers in each test. However, deletion of all these would have resulted in the loss of 48 cases, and the possible loss of relevant information for an adaptation study. Inspection of the P–P graphs for departures of normality suggested that observed AWMA test scores were approximately normal, except for Mr. X. This latter test was subjected to a log transformation, which did not have the desired normalizing effect. Therefore, with these considerations in mind, we performed the analyses with parametric as well as non-parametric approaches, when possible. As it will be reported, both of them had similar results.

1.3.3.1. Item analyses. Difficulty level was studied as proportion of correct items. Fig. 3 shows an increase in difficulty as the number of to-be-remembered items was raised.

1.3.3.2. Reliability. Internal consistency (Cronbach's α) was calculated for each of the AWMA tests. Table 5 shows high values for all of the tasks.

1.3.3.3. Validity. Table 6 shows Pearson's r correlations between all AWMA tests, which were of medium to high size (between .31 and .79), and significant ($p < .001$, adjusting the family-wise error to the number of contrasts with the Bonferroni correction).

Table 4
Descriptive statistics for AWMA subtests and cognitive tests.

	N	Median	Minimum	Maximum	Mean	Standard deviation	Skewness	Kurtosis
Digit Recall – AWMA	210	23.00	13	42	23.43	4.69	.609	1.023
Word Recall – AWMA	210	18.00	0	30	17.42	4.72	−.105	.183
Nonword Recall – AWMA	210	9.00	0	21	9.48	3.44	.247	.061
Dot Matrix – AWMA	210	21.00	7	36	20.79	5.95	.184	−.208
Block Recall – AWMA	210	20.00	0	35	18.80	8.13	−.593	−.374
Mazes Memory – AWMA	210	20.00	6	36	20.01	6.42	.005	−.502
Listening Recall – AWMA	210	6.50	0	23	7.78	5.16	.425	−.053
Counting Recall – AWMA	210	16.00	6	29	15.60	5.46	.105	−.640
Backward Digit Recall – AWMA	210	9.00	0	29	9.89	4.61	.567	.902
Odd One Out – AWMA	210	15.00	0	30	14.63	5.75	.090	−.655
Mr. X – AWMA	210	6.00	0	24	7.72	4.76	.737	.357
Spatial Span – AWMA	210	14.00	0	29	13.40	6.26	−.112	−.799
Tower of London	210	22.00	2	56	24.51	11.71	.418	−.862
Picture Span	210	60.00	2	84	57.25	17.83	−1.219	1.420
Word Order	210	9.00	1	13	8.83	2.10	−.241	.077
Block Design	210	9.00	2	15	8.74	2.64	.131	−.537

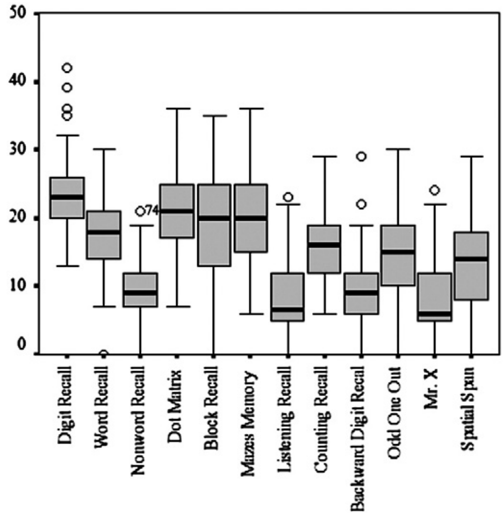

Fig. 2. Scores distribution for AWMA tests.

[Table 7](#) shows correlations between AWMA scores and Picture Span (PS), Word Order (WO), Tower of London (TOL), and Block Design (BD). Regarding convergent validity, positive and significant correlations were found among the AWMA tests, two tests of working memory (PS and WO), and a test of executive functioning (TOL), as seen in Table 7. On the other hand, with respect to discriminant validity, correlations between the AWMA tests and Block Design were low and non-significant.

2. General discussion

The adapted AWMA was well understood by all children in both studies (pilot and adaptation). The three verbal tests adapted, as well as the other tests, showed a steady increase in difficulty as a function of item size, or to-be-remembered items. These considerations suggest that administration (instructions and item order and presentation) of the adapted battery has an adequate functioning.

Distribution of scores in some tests (Digits, Backward Digits, Listening Recall, and Mr. X) did not fit the normality assumption. These results are probably due to the lack of smaller and older aged groups of children, and also to the exclusion criteria. If so, further studies including other age groups, and children with school problems, or with psychiatric or neurologic disabilities, might solve this issue.

Regarding reliability, a high internal consistency (Cronbach's α) was observed for all the tests. Alloway et al. (2006) assessed the AWMA reliability using the test-retest method; therefore reliability across studies cannot be compared.

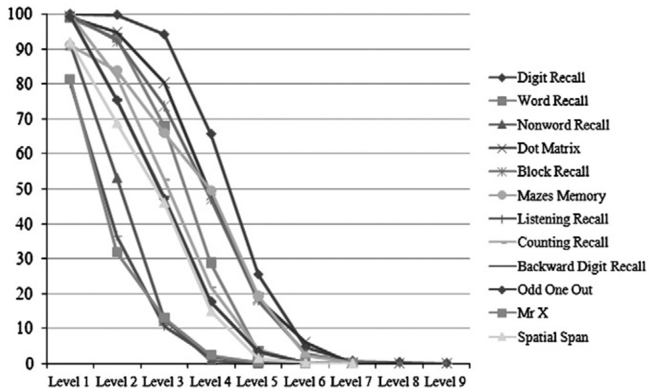

Fig. 3. Percent correct responses per trial size in the AWMA tests.

Table 5
Cronbach's α for each AWMA subtest.

	Alpha coefficient
Digit Recall	.88
Word Recall	.87
Nonword Recall	.81
Dot Matrix	.91
Block Recall	.92
Mazes Memory	.95
Listening Recall	.91
Counting Recall	.90
Backward Digit Recall	.89
Odd One Out	.91
Mr. X	.89
Spatial Span	.92

Correlations among all the tests were from moderate to high (between .31 and .79), and all significant. This pattern supports the claim that all tests were tapping the same, or closely related, constructs. These findings are consistent with the results presented by Alloway et al. (2006), where all correlations were moderate to high and significant, between .46 and .74. Correlations of similar value were found in our sample and Alloway et al. (2006). The only exception was that our study obtained lower correlations between Word Recall and the following tests (first r: this study; second r: Alloway et al, 2006): Digit Recall ($r = .36$ vs. $r = .64$), Nonword Recall ($r = .31$ vs. $r = .73$), Dot Matrix ($r = .33$ vs. $r = .58$), Mazes Memory ($r = .35$ vs. $r = .60$), Listening Recall ($r = .32$ vs. $r = .60$), Counting Recall ($r = .37$ vs. $r = .63$), Backward Digit Recall ($r = .34$ vs. $r = .60$), Mr. X ($r = .36$ vs. $r = .50$), and Spatial Span ($r = .33$ vs. $r = .55$); and Nonword Recall with two tests: Dot Matrix ($r = .34$ vs. $r = .51$) and Block Recall ($r = .29$ vs. $r = .50$). Thus, in general the correlations among tests are similar in both versions except for Word Recall, which showed systematic lower correlations with all tests. This could be due to differences in processing or factorial structure in both samples, a question worth exploring in future studies.

As for convergent and discriminant validity, moderate (.23 to .64) and significant correlations with other measures of working memory (Picture Span, Word Order) and executive function (Tower of London), strengthen the validity of the AWMA measures as tapping aspects of working memory functioning. Non-significant correlations with Block Design, a measure of visuo-constructive ability, even for visuospatial measures of WM (as in previous studies, e.g. Luciana et al, 2005) add to the discriminant validity of the AWMA.

We compared the results from this study with those published by Alloway et al. (2006). Under a .05 family-wise error adjusted for all contrasts with the Bonferroni correction, we found significant differences in Digit Recall ($t_{(916)} = 6.140$; $p < .001$), Listening Recall ($t_{(916)} = 4.471$; $P < .001$), Nonword Recall ($t_{(916)} = 6.359$; $p < .001$), and Mr. X. ($t_{(916)} = 4.071$; $p < .001$) in favor of the UK sample and in Dot Matrix ($t_{(916)} = 5.648$; $p < .001$), Mazes ($t_{(916)} = 8.446$; $p < .001$), and Block Recall ($t_{(916)} = 5.518$; $p < .001$) in favor of the Argentinean sample. The Spanish-speaking sample had lower means in three verbal and one visuospatial test. On the other hand, the UK sample had lower scores on three visuospatial tests. Whether these results are due to ability or processing differences in both samples, or to test artifacts, would be a question for further exploration.

Since working memory has an important role during childhood in different cognitive abilities, as reading, language comprehension, vocabulary acquisition, arithmetic, among others, it is crucial to be able to have an instrument that allows discriminating children with working memory impairment and eventually treating them. On the clinical level, is also important to have an instrument that could complete and complement the diagnosis and treatment of a number of pathologies as Developmental Coordination Disorder, Specific Language Impairment, Attention-Deficit/Hyperactivity Disorder, and Asperger syndrome.

Table 6
Pearson's correlations among AWMA subtests.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Digit Recall	.	.36**	.54**	.51**	.56**	.54**	.66**	.59**	.62**	.52**	.42**	.48**
2. Word Recall		.	.31**	.33**	.40**	.35**	.32**	.37**	.34**	.39**	.36**	.33**
3. Nonword Recall			.	.34**	.29**	.41**	.48**	.41**	.44**	.41**	.38**	.36**
4. Dot Matrix					.68**	.79**	.57**	.67**	.59**	.71**	.58**	.62**
5. Block Recall						.69**	.64**	.65**	.63**	.64**	.60**	.63**
6. Mazes Memory							.63**	.75**	.65**	.72**	.61**	.66**
7. Listening Recall								.69**	.69**	.61**	.58**	.58**
8. Counting Recall									.70**	.70**	.53**	.63**
9. Backward Digit Recall										.65**	.55**	.62**
10. Odd One Out											.63**	.65**
11. Mr. X												.59**
12. Spatial Span												

Note. $N = 210$.
**$p < .01$ adjusting the family-wise error to the number of contrasts with the Bonferroni correction.

Table 7
Correlations between AWMA subtests and cognitive tests.

	TOL[a]	PS[b]	WO[c]	BD[d]
	r^e	r^e	r^e	r^e
Digit Recall	.23**	.35**	.64**	−.13
Word Recall	.24**	.24**	.33**	.00
Nonword Recall	.29**	.24**	.44**	.00
Dot Matrix	.41**	.32**	.39**	−.02
Block Recall	.38**	.34**	.45**	−.06
Mazes Memory	.42**	.40**	.42**	−.07
Listening Recall	.46**	.38**	.49**	−.05
Counting Recall	.38**	.38**	.45**	−.02
Backward Digit Recall	.41**	.37**	.53**	−.03
Odd One Out	.38**	.39**	.38**	−.03
Mr. X	.41**	.38**	.30**	−.03
Spatial Span	.34**	.40**	.41**	−.03

Note. $N = 210$; a: Tower of London; b: Picture Span; c: Word Order; d: Block Design; e: Pearson's r.
**$p < .01$ adjusting the family-wise error to the number of contrasts with the Bonferroni correction.

Overall, the study shows that the adapted AWMA can be considered a valid and reliable battery of working memory, suitable to be employed with Spanish-speaking children from 6 to 11 years old in Argentina. Further research is needed to validate it in other age groups and Spanish speaking communities.

Acknowledgments

This research was supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (Res. No 258/06 and Res. No 3100/08) and by the Secretaría de Ciencia y Técnica, Universidad de Buenos Aires (UBACyT P016).

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30−60.

Alloway, T. P. (2007a). *Automated Working Memory Assessment.* London: The Psychological Corporation.

Alloway, T. P. (2007b). Working memory, reading, and mathematical skills in children with developmental coordination disorder. *Journal of Experimental Child Psychology*, *96*, 20−36.

Alloway, T. P., & Alloway, R. G. (2009). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20−29.

Alloway, T. P., & Archibald, L. M. D. (2008). Working memory and learning in children with developmental coordination disorder and specific language impairment. *Journal of Learning Disabilities*, *41*(3), 251−262.

Alloway, T. P., Gathercole, S. E., Adams, A. M., Willis, C., Eaglen, R., & Lamont, E. (2005). Working memory and other cognitive skills as predictors of progress toward early learning goals at school entry. *British Journal of Developmental Psychology*, *23*, 417−426.

Alloway, T. P., Gathercole, S. E., Kirkwood, H. J., & Elliot, J. G. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*(2), 606−621.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the Automated Working Memory Assessment. *Journal of Educational Psychology*, *28*(7), 725−734.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*(6), 1698−1716.

Alloway, T. P., Rajendran, G., & Archibald, L. M. D. (2009). Working memory in children with developmental disorders. *Journal of Learning Disabilities*, *42*(4), 372−882.

Baddeley, A. D. (1986). *Working memory.* Oxford: Clarendon Press.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, *49A*, 5−28.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158−173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake, & P. Shah (Eds.), *Models of Working Memory*. Cambridge: Cambridge University Press.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, task switching, and working memory. *Developmental Neuropsychology*, *19*, 273−293.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral & Brain Sciences*, *22*, 77−126.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769−786.

De Jong, P. F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology*, *70*, 75−96.

Dixon, P., LeFevre, J. A., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*, *80*(4), 465−472.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and the functions of the prefrontal cortex. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY: Cambridge University Press.

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory and Cognition*, *28*(5), 774−782.

Gathercole, S. E., Alloway, T. P., Kirkwood, H. J., Elliot, J. G., Holmes, J., & Hilton, K. A. (2008). Attentional and executive function behaviours in children with poor working memory. *Learning and Individual Differences*, *18*(2), 214−223.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language.* Hove: Earlbaum.

Gathercole, S. E., Brown, L., & Pickering, S. J. (2003). Working memory assessments at school entry as longitudinal predictors of national curriculum attainment levels. *Educational and Child Psychology*, *20*, 109−122.

Gathercole, S. E., & Pickering, S. J. (2000). Assessment of working memory in 6- and 7-year-old children. *Journal of Educational Psychology*, *92*, 377−390.

Hale, S., & Fry, A. F. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*, 1−34.

Injoque-Ricle, I., & Burin, D. (2008a). Validación preliminar de una prueba de memoria de trabajo para niños: Serie de Dibujos. *Memorias de XV Jornadas de Investigación en Psicología, Tomo II* (pp. 331−332).

Injoque-Ricle, I., & Burin, D. I. (2008b). Validez y fiabilidad de la prueba de Torre de Londres para niños. *Revista Argentina de Neuropsicología*, *11*, 21−31.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122−149.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working-memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 21−48). Oxford University Press.

Kaufman, A. S. (1983). *Batería de evaluación de Kaufmann para niños*. España TEA Ed.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children* (2nd ed.). Circle Pines, MN: American Guidance Service K-ABC-II.

Long, D. L., & Prat, C. S. (2008). Individual differences in syntactic ambiguity resolution: Readers vary in their use of plausibility information. *Memory and Cognition*, *36*, 375—391.

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal and executive control processes in adolescents. *Child Development*, *76*, 697—712.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*, 35—54.

Miyake, A., & Shah, P. (1999). Toward unified theories of working memory. Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake, & P. Shah (Eds.), *Models of working memory. Mechanisms of active maintenance and executive control*. New York, NY: Cambridge University Press.

Pickering, S. J., & Gathercole, S. E. (2004). Distinctive working memory profiles in children with special educational needs. *Educational Psychology*, *24*, 393—408.

Real Academia Española (2001). *Diccionario de la real academia española, 22° Ed.* Madrid: Espasa-Calpe.

Sebastián Galles, N., Martí Antonín, M. A., Carreiras Valiña, M. F., & Cuetos Vega, F. (2000). *LEXESP: Léxico informatizado del Español.* Barcelona: Edicions de la Universitat de Barcelona.

Shallice, T. (1982). Specific impairments in planning. *Philosophical Transcripts of the Royal Society of London*, *298*, 199—209.

Vallar, G., & Shallice, T. (1990). *Neuropsychological impairments of short-term memory*. Cambridge: Cambridge University Press.

Wechsler, D. (1994). *Test de Inteligencia para niños WISC-III.* Buenos Aires: Paidós.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children manual* (4th ed.). San Antonio, TX: Psychological Corporation.