JIDR Journal of Intellectual Disability Research
Published on behalf of mencap and in association with IASSID

Journal of Intellectual Disability Research doi: 10.1111/j.1365-2788.2010.01285.x

VOLUME 54 PART 5 pp 433–447 MAY 2010

Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities

M. J. Van der Molen,[1] J. E. H. Van Luit,[1] M. W. Van der Molen,[2] I. Klugkist[3] & M. J. Jongmans[1,4]

1 *Utrecht University, Department of General and Special Education, Utrecht, the Netherlands*
2 *University of Amsterdam, Department of Psychology, Amsterdam, the Netherlands*
3 *Utrecht University, Department of Methodology & Statistics, Utrecht, the Netherlands*
4 *Department of Pediatric Psychology, Wilhelmina Children's Hospital, University Medical Centre Utrecht, the Netherlands*

Abstract

Background The goal of this study is to evaluate the effectiveness of a computerised working memory (WM) training on memory, response inhibition, fluid intelligence, scholastic abilities and the recall of stories in adolescents with mild to borderline intellectual disabilities attending special education.
Method A total of 95 adolescents with mild to borderline intellectual disabilities were randomly assigned to either a training adaptive to each child's progress in WM, a non-adaptive WM training, or to a control group.
Results Verbal short-term memory (STM) improved significantly from pre- to post-testing in the group who received the adaptive training compared with the control group. The beneficial effect on verbal STM was maintained at follow-up and other effects became clear at that time as well. Both the adaptive and non-adaptive WM training led to higher scores at follow-up than at post-intervention on visual STM, arithmetic and story recall compared with the control condition. In addition, the non-adaptive training group showed a significant increase in visuo-spatial WM capacity.
Conclusion The current study provides the first demonstration that WM can be effectively trained in adolescents with mild to borderline intellectual disabilities.

Keywords cognitive intervention, remediation, lower intellectual functioning, randomized control trial

Correspondence: Dr Mariët Van der Molen, University of Amsterdam, Department of Psychology, Roetersstraat 15, 1018 WB, Amsterdam, the Netherlands (e-mail: M.J.vanderMolen@UvA.nl).

Introduction

Working memory (WM), the ability to maintain and process information simultaneously during the performance of a cognitive task, is central to fluid intelligence (e.g. Ackerman *et al.* 2005; Friedman *et al.* 2006; Jaeggi *et al.* 2009) and executive functioning (e.g. Garon *et al.* 2008), and is critical to the development of children's scholastic abilities such as arithmetic and reading (e.g. Hitch *et al.*

2001; Alloway *et al.* 2009). It has been suggested that even a small increase in the efficacy of WM will lead to significant improvements in classroom and daily life functioning in children (Minear & Shah 2006). WM training studies have been shown effective in various populations including the elderly (Craik *et al.* 2007), people with schizophrenia (Kurtz *et al.* 2007), typically developing children with low WM capacity (Holmes *et al.* 2009), children with attention deficit/hyperactivity disorder (ADHD; Klingberg *et al.* 2005) and children with traumatic brain injuries (Van 't Hooft *et al.* 2007).

The focus of the current study is on adolescents with mild to borderline intellectual disabilities (M-BID; IQ score 55–85). This group is known to have substantial WM problems (Henry 2001; Van der Molen *et al.* 2007, 2009; Maehler & Schuchardt 2009), generally performs poorly on academic achievement domains (Sabornie *et al.* 2005) and requires more educational support than do typically developing adolescents (Simonoff *et al.* 2006). Given the relationship between WM performance and scholastic abilities, it is of substantial interest to study the feasibility and effectiveness of a WM training in adolescents with M-BID.

Working memory is considered to depend simultaneously on (verbal or visuo-spatial) short-term memory (STM) and the control and regulation of attention (or the so-called 'central executive' system; Baddeley 1986; Cowan 1999; Engle RE *et al.* 1999). Information held in verbal STM can be rehearsed automatically by typically developing individuals with a chronological age of 7 and older (Henry & Miller 1993) to prevent the information from fading away. Studies investigating the effects of memory training programmes in people with intellectual disabilities (ID) focused primarily on verbal STM; more specifically on the rehearsal aspect of STM (e.g. Kellas *et al.* 1973; Engle & Nagle 1979; Comblain 1994; Laws *et al.* 1996; Conners *et al.* 2001, 2008). Broadly speaking, these training programmes proved beneficial, in that individuals with ID showed gains in their ability to repeat items in the correct order. Unfortunately, however, spontaneous application of the rehearsing strategy beyond the training domain appeared problematic (e.g. Bebko & Luhaorg 1998). Moreover, follow-up studies showed that the effect of this specific training deteriorates with time (e.g. Laws *et al.* 1995).

In contrast to verbal STM, individuals with ID have not been presented with WM training (Minear & Shah 2006). This omission is unfortunate because, as discussed above, WM is especially weak in children and adolescents with M-BID and it is WM above all that is associated with scholastic abilities (e.g. Alloway *et al.* 2009). Moreover, a computerised WM training yielded positive results in children with ADHD (Klingberg *et al.* 2005; see also Klingberg *et al.* 2002, for a study in small groups of children with ADHD and university students) and in typically developing children with low WM capacity (Holmes *et al.* 2009). Klingberg *et al.* (2005) demonstrated that in children with ADHD, it is possible to improve visual and verbal WM, response inhibition and fluid intelligence. The effects were shown immediately after training and at follow-up, although at follow-up the effect on fluid intelligence disappeared. Holmes *et al.* (2009) showed that typically developing children with low WM capacity profited from the same training as expressed in higher scores on visuo-spatial STM, verbal WM, visuo-spatial WM and a practically based assessment of WM use in the classroom immediately and 6 months after the training. Furthermore, they observed beneficial effects on one of the four administered ability measures, a mathematical reasoning test at follow-up testing.

In the current study, a WM training was used developed analogous to the Cogmed Working Memory Training devised by Klingberg and co-workers (e.g. Klingberg *et al.* 2005). Our training, coined 'Odd Yellow', was based on our reading of the pertinent literature and on recommendations of an expert panel of professionals with extensive experience in working with children and adolescents with M-BID. The 'Odd Yellow' is based on the WM task 'Odd-One-Out', as included in the 'Automated Working Memory Assessment' (Alloway 2007), used in studies in children and adolescents with M-BID (Henry 2001; Van der Molen *et al.* 2007). In the 'Odd Yellow', a sequence of three similar looking figures is shown on the computer screen. Two of the figures are identical in shape and the other is slightly different, the 'odd-one-out'. The figures are drawn in black, except one of the two identical shapes, which is yellow (see also Fig. 2). For every sequence, the child has to reproduce the location of the odd-one-out and the location of the yellow figure shape. The training starts with trials consisting of one sequence and the number of sequences increases with training up to seven sequences. During the course of training, the number of sequences is dynamically adjusted to the child's performance level. When the child starts making mistakes, the number of sequences decreases to the level the child can handle at that moment. The duration of each training session, 6 min, was well within the attention span of most adolescents with M-BID as observed in our pilot study. The computer screen included a time counter for encouraging the child to complete the session. In addition, the child was told that the completion of the training would be rewarded with a personal certificate. The training consisted of 15 sessions across 5 weeks, with a minimum of two and a maximum of four sessions per week. Finally, the computerised training allowed for evaluating training compliance and data transmission using shielded Internet connections.

Three versions of the 'Odd Yellow' WM training were constructed: Training A used a tracking procedure to adjust level of difficulty to the individual's momentary capacity. It is a demanding training as it challenges continuously the WM capacity of the children during each session. This adaptive aspect is seen in most other WM training programmes and has proven to be successful in for example children with ADHD (e.g. Klingberg *et al.* 2005) and in typically developing children with low WM skills (Holmes *et al.* 2009). Training B used a fixed and low level of difficulty aiming at presenting the same training routine but without challenging WM capacity. This version was included as we wanted to explore if stimulating WM alone, without challenging it, would boost these children's weak cognitive component. The control training finally, was nearly identical to training A but without placing any demand on memory capacity.

To evaluate the effect of the training, an extensive test battery was administered immediately before and following the training, as well as 10 weeks after the training. As the training is targeting the ability to both simultaneously process and store information at the short term, we consider WM and STM tasks as the primary outcome measures. Previous studies have shown effects on STM and WM measures immediately after WM training and at follow-up (e.g. Klingberg *et al.* 2005; Holmes *et al.* 2009). Therefore, it is expected that both children with M-BID in training groups A and B perform better at post-testing and at follow-up compared with the children in the control group on the primary outcome measures. Secondary outcome measures focus on scholastic and cognitive aspects that are said to be influenced by WM capacity. More specifically, studies showed that WM training has a beneficial effect on response inhibition and fluid intelligence immediately after the training (Klingberg *et al.* 2005) and on mathematical reasoning at follow-up assessment (Holmes *et al.* 2009). Therefore, we included measurements of these three aspects in the current study. Furthermore, we included two more tasks measuring aspects that are said to be influenced by WM capacity; reading ability (e.g. Gathercole & Pickering 2000) and story recall (Van der Molen *et al.* in press). Based on the aforementioned studies, it is expected that the children in training groups A and B will perform better on the secondary outcome measures after the training than the control group.

Method

Participants

Seven special schools for children with M-BID participated in the study. A criterion for entrance in this type of school is an IQ score in the range 55–85. The schools were located in three different provinces in the Netherlands. In total, 95 adolescents within the age range of 13–16 years were selected to participate in the training (mean age 15.21 years and SD = 0.69). Adolescents with a DSM-IV-TR (American Psychiatric Association 2000) diagnosis of ADHD or Autism Spectrum Disorder (ASD) were excluded from participation, as these psychiatric problems are known to be associated with specific WM strengths and weaknesses (Minear & Shah 2006). Adolescents who were previously hospitalised because of serious (head) injuries were also excluded. From two adolescents there were no scores from the post-tests (because of illness) and for six other adolescents there were no scores from the follow-up tests (all because of illness). See Fig. 1 for the flow of participants.

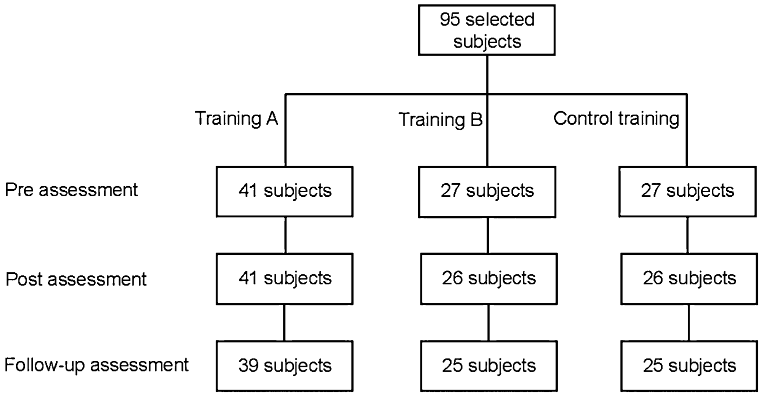


Figure 1 Flow of participants through the trial.

The adolescents were randomly assigned to either training A (group A), training B (group B) or to the control training (control group). Group A consisted of 41 adolescents (23 boys, 18 girls) with a mean age of 15.32 years (SD = 0.68) and, as an index of fluid intelligence, a mean Raven score of 35.37 (SD = 6.26), group B contained 27 adolescents (15 boys, 12 girls) with a mean age of 15.00 years (SD = 0.70) and a mean Raven score of 32.70 (SD = 7.91) and the control group also included 27 adolescents (16 boys, 11 girls), with a mean age of 15.43 years (SD = 0.66), and a mean Raven score of 33.37 (SD = 5.29). The Raven scores of the three groups fell between the fourth and fifth percentile. The three groups did not differ in age, $F_{2,92} = 2.2$, $P = 0.12$, nor in fluid intelligence (Raven), $F_{2,92} = 1.4$, $P = 0.24$. Ethnicity and social economical status were comparable over all three groups.

Study design

A randomised, single-blind controlled trial was conducted. Schools were approached and informed about the study and the inclusion and exclusion criteria. After obtaining informed consent, all adolescents were tested five times with the same, single version, test battery. They were tested twice using the same test battery within a 4-week period (pre-test 1 in the first 2 weeks and pre-test 2 in the last 2 weeks). This was followed by a 5-week training programme. Following the training, the adolescents were tested twice again within a 4-week period (post-test 1 in the first 2 weeks and post-test 2 in the last 2 weeks). Finally, the test battery was administered for the last time (follow-up test) just before the summer holiday, which was 10 weeks after the training was finished. The double-testing at pre- and post-assessment was to avoid the potential contribution of negative factors like fatigue, influencing the outcomes. Therefore, the double scores of each test were averaged. However, the Raven Standard Progressive Matrices (SPM, Raven *et al.* 1996) was only administered ones at pre- and post period. Furthermore, during follow-up assessment, all measurements were only single-tested in order to prevent the children of becoming bored with the assessments. Thus all tests (except the Raven SPM) were administered five times. The current assessment procedure carried the obvious danger of test–retest confounds. It was anticipated, however, that test–retest effects are similar for the three groups and, thus, will not contaminate any training effects in between-groups comparisons.

Within schools, the assignment of each adolescent to one of the three conditions (i.e. training A, training B, or the control training) was blind and at random. As we anticipated training A to be the most effective, we assigned a relatively large group of children to that training condition. Teachers prompted the adolescents to perform the training sessions at the specified times (three times a week). The computers used for training (both PCs and Macs were used) were connected to the Internet and the automatically sent data were checked by the experimenters for training compliance and quality. After follow-up testing, all adolescents received a small present and a certificate testifying they had completed a memory training.

Primary outcome measures

STM and WM

Two verbal and two visual STM tests were used. Digit Recall and Nonword Recall (Pickering & Gathercole 2001) both measure verbal STM. Both tests require repeating digits or nonwords in the same order as presented. Digit Recall starts with two digits up to eight, while Nonword Recall starts with one nonword up to six. For these, and all of the following span measures (except the Visual Patterns test), there are six trials per list length. List lengths increase incrementally on the condition that at least four of the six trials are completely correct, the omitted trials are awarded one point each. Memory scores represent the number of trials that were completely correct. Scores vary from 0 to 42 (digits) or from 0 to 36 (nonwords).

Visual STM was assessed by using Block Recall and the Visual Patterns test. Block Recall is identical to the Corsi test (see Lezak, 1995), but in this study we used the instructions from Pickering & Gathercole (2001). The experimenter taps a sequence of three-dimensional blocks that the child has to repeat in the same order. The task starts with one block up to sequences of nine blocks. Scores vary from 0 to 54. In the Visual Patterns test (Della Sala *et al.* 1997), the child is shown a matrix depicted on a stimulus card, varying from 2×2 to 5×6 squares with half of the squares being marked. After inspecting a stimulus card for 3 s, the child has to indicate the marked squares using a blank grid on the response sheet. Three stimulus cards are available for each of the 14 difficulty levels. List length increases incrementally on the condition that at least two of the three trials are completely correct. Scores vary from 0 to 42.

Two verbal and one visuo-spatial WM tests were used. The two verbal WM tests were Backward Digit Recall and Listening Recall. Backward Digit Recall (Pickering & Gathercole 2001) requires repeating spoken lists of digits, but in the reverse order. Listening Recall (Pickering & Gathercole 2001) requires listening to simple statements to determine whether they are true or false, while at the same time remembering the last word of each statement. Following each trial, these last words are to be repeated in the same order as presented. Trials in Backward Digit Recall start with two digits up to seven, while Listening Recall starts with one sentence, up to a maximum of six. Scores vary from 0 to 36 for each of these tests.

Visuo-spatial WM was examined using a manual version of the Spatial Span (Alloway 2007). A card is shown with two shapes of which the right one has a red dot on top. The right shape can be exactly the same (p–p) or opposite (p–q) to the left shape and it can be rotated in three different ways (0°, 120° and 240°). The child has to decide whether the shape at the right is the same or opposite to the left shape. At the same time, the position of the red dot on the right shape has to be remembered, which can be at three different locations according to the three rotation possibilities. After each trial, the child has to point to one of three dots (at 0°, 120° or 240°) to indicate which dots were on the stimuli cards and in which sequence. Trials start with one card up to a sequence of six. Scores can vary from 0 to 42.

Secondary outcome measures

Scholastic abilities

Two tests were administered to tap scholastic abilities, one test for arithmetic and another one for reading abilities. The Arithmetic test (De Vos 1992) presents the child with five rows for different arithmetic operations: adding, subtracting, multiplying, dividing and a row combining the four operations. The child has to complete as many items in each row as possible within 1 min by writing down the correct answer. For every correct answer, one point is given. Total score is the total amount of correct scores for all five rows with a minimum of 0 and a maximum of 200. The Reading test (Brus & Voeten 1973) presents the child with a list containing 116 unrelated words of increasing difficulty. The child has to read aloud as many words as possible within 1 min. Total score is the amount of correct read words varying from 0 to a maximum of 116.

Story Recall

In Story Recall (Van der Molen 2007), the experimenter reads out loud a short story after which the child immediately has to repeat it as exactly as possible (immediate recall). After 20 min, the child is asked to repeat the story a second time (delayed recall). Total score, indexed as correct key words of the story in both immediate and recall condition, varies from 0 to 29.

Response inhibition

The Stroop (Hammes 1978), measuring response inhibition, consists of three cards. First, the participant has to read as quickly as possible the names of four colours (yellow, red, green and blue) written on the first card. Then on the second card, the participant sees blocks filled in with the four colours and has to name these as quickly as possible. Finally on the third card, the words of the four colours are written and printed in a different colour. The participant has to name the colour in which the words are printed and inhibit the prepotent response to name the word. The total (inference) score is the amount of seconds needed to read out the third card minus the seconds needed for the second card. Scores can vary as they depend on how long the participant takes to read aloud card two and card three. The less interference time, the better the achievement.

Fluid intelligence

Finally, the Raven SPM (Raven *et al.* 1996) was administered to assess fluid intelligence. The participant is presented with pieces of wallpaper on which one part is missing. The participant has to choose the correct missing part out of different alternatives. Other tasks within this test consist of rows of symbols with a missing one: the participant has to decide which of the alternative symbols given, logically completes the row. The test starts relatively easy but increases in difficulty. One point is given for each correct answer. Minimum score is 0 and maximum score is 60.

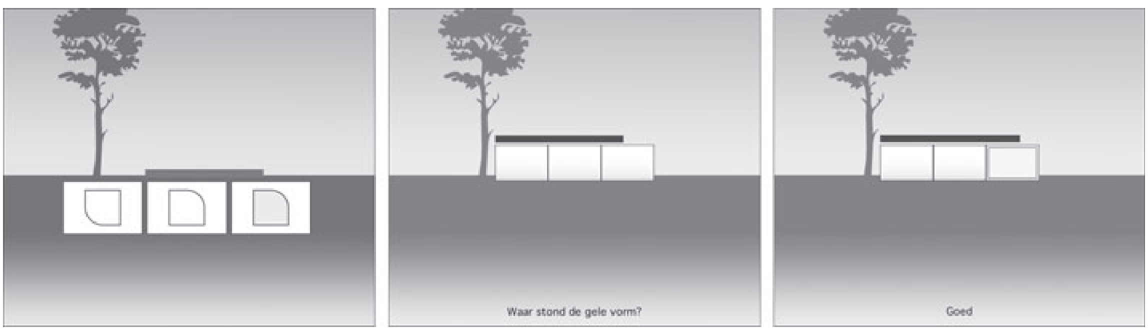
Motivation

Likert scales (1–10) were used to assess how much the participants liked their training and how much effort they invested in it.

Training programme

Training A

We developed a computerised 'Odd Yellow' training based on the principle of the Odd-One-Out test by Henry (2001). Three times a week, during a 5-week period, the participant is trained for 6 min. After logging in, the session starts with a brief explanation. Then the participant is shown three figures (for an example see Fig. 2), two of which are identical in shape and one is slightly different in shape, the odd-one-out. Also, of the three figures, two are drawn in black and one in yellow (which cannot be the odd-one-out). First, the participant has to tap the odd-one-out as quickly as possible by use of the mouse prompted by a visual timer, a line which



(a) (b) What was the location of the grey figure? (c) Correct

Figure 2 Stills from the working memory training. Three figures are shown. The figure, which is slightly different in shape, in this case the left one, has to be clicked on with the mouse. The grey line on top of the symbols, i.e. the time bar, fades to the left so as to force the child to choose the odd-one-out within 5 s (a). Then the child gets 2 s to remember the place of the grey figure, after which the child is asked to click in the empty box where the grey figure went (b). The child clicked in the right box (c).

decreases in length and disappears within 5 s (see Fig. 2). After clicking the odd-one-out, or following 5 s when the participant omitted to respond (receiving the feedback 'too late'), two seconds are provided to remember the position of the yellow figure. These two seconds are represented by a visual timer as well. When the two seconds are expired, an empty matrix is shown representing the positions of the three figures. The participant, again forced by a 5-s visual timer, has to tap the position of the yellow figure. Trials start with one sequence of three figures and can increase to seven sequences of three figures each. For example, in trials of two sequences, the participant clicks on the odd-one-out in sequence 1, then clicks on the odd-one-out in sequence 2, and then is prompted to identify in an empty matrix representing the two sequences, the position of the yellow figure in the first sequence and then the position of the yellow figure in the second sequence. When two trials of the same sequence are failed, that is when the positions of the yellow figures are not remembered correctly, trials with one sequence less are presented.

Training B

This training is similar to training A except that trials do not exceed more than two sequences. The participant is presented with trials of one or of two sequences in random order.

Control training

The format for the control training is similar to training A, except that now only the odd-one-out has to be detected as soon as possible while the position of the yellow figure can be ignored. That is, the participant has to tap the odd-one-out in each sequence as quickly as possible by use of the mouse prompted by a visual timer, exactly like in both training conditions. However, in the control condition, the participant is offered sequences of three figures requiring only the identification of the odd-one-out.

Data analysis

First, the motivation rates were compared between the three groups by use of MANOVA. Then data from both pre-test sessions (pre-test 1 and pre-test 2) and post-test sessions (post-test 1 and post-test 2) were averaged resulting in three scores per participant: mean pre-test score, mean post-test score and one follow-up test score.

As the interest of the training effects centred on constructs rather than single test scores, the analyses were done on combined test scores indexing the same underlying construct. Because the tests had different scales, individual test scores all were first linearly transformed to a scale with a minimum score of 0 and a maximum score of 10. An exception was made for the Stroop test as it has no maximum score, so the original score in seconds was retained. This resulted in eight variables: verbal STM (compound score for Digit Recall and Nonword Recall), visual STM (Block Recall and Visual Patterns test), verbal WM (Listening Recall and Backward Digit Recall), visuo-spatial WM (Spatial Span), scholastic abilities (arithmetic test and reading test), Story Recall (immediate and delayed), response inhibition (Stroop) and fluid intelligence. Unlike the other measures, a lower score on the Stroop indicates a better performance.

We contrasted the performances of both groups A and B with the control group. As we expected both groups A and B to perform better than the control group after the training, planned comparisons were used (i.e. one-sided). Similar to Klingberg *et al.* (2005), a general linear model analysis (GLM) was performed controlling for baseline scores (pre-testing). To assess whether the observed effects were stable, the outcome measures at follow-up were examined using a GLM, controlling for post-testing score. When significant effects were found for a compound variable, the individual test scores, contributing to this particular compound score, were analysed, so as to see if both or one of the test scores contributed to the effect.

Results

Groups did not differ in terms of how much they liked the training, with mean scores varying between 7.2 and 7.5, or how much effort they invested in the training, with mean scores varying from 7.8 to 8.1 ($F_{2,91} = 0.35$, $P = 0.85$).

During the training, the mean maximum training span of training group A increased till around session 9 (see Fig. 3).

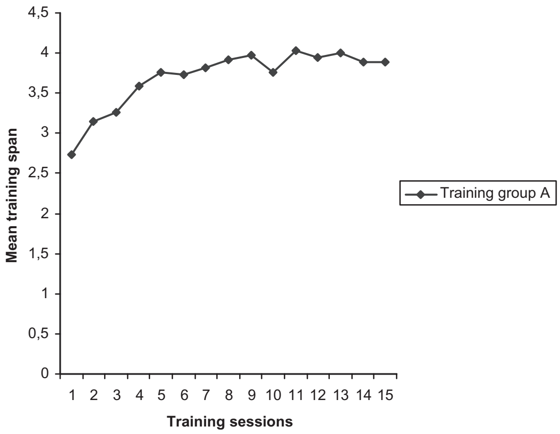


Figure 3 Mean maximum training span for training group A.

Mean raw scores (SD) of the three groups on all tests are shown in Table 1a,b.

The outcome of the GLMs revealed that group A obtained a significant higher score on the verbal STM compound score in the post-training condition than at pre-training condition compared with the control group (see Table 2 for η^2_p, B, the non-standardised regression coefficient, and one-sided P for all compound scores and for mean values corrected for baseline measurement. The progress in scores from pre- (time 1) to post- (time 2) to follow-up testing (time 3) for each of the three groups is shown in Fig. 4). Analyses of the individual test scores showed that the mean Digit Recall score at this point in time was higher for group A than for the control group ($\eta^2_p = 0.04$, $B = 0.94$, $P = 0.03$), while this was not the case for scores on Nonword Recall. Between post-testing and follow-up no significant change was observed, indicating that the positive effect on verbal STM remained at follow-up.

No other significant results were found from pre- to post-testing for group A. Group B did not show any significant change between pre- and post-testing.

From post- to follow-up assessment significant differences appeared on several scores for both group A and group B compared with the control group. Compound Visual STM scores increased for group A and for group B compared with the control group. Analyses on the individual visual STM test scores indicated that for both groups the increase was explained by a higher score on Block Recall. Visuo-spatial WM, indexed by the task Spatial Span, also increased from post- to follow-up test for group B. Scholastic abilities compound score increased for both group A and group B compared with the control group. In both cases, it was an increase in score on the Arithmetic test and not on the Reading test, which was responsible for the change. Finally, Story Recall compound score was higher (i.e. better) at follow-up than at post-testing for group A and for group B. Increases in both Story Recall Immediate and Story Recall Delayed were responsible for the change.

The increases in the different compound scores from post- to follow-up tests were also examined vis-à-vis the base-line (pre-testing). GLM analyses were run on follow-up scores with pre-test score as covariate. Group A had a higher visuo-spatial WM score (Spatial Span), compared with the control group. Group A and group B obtained a higher compound score on Scholastic abilities, caused by a higher score on the Arithmetic test. Group A and group B did also obtain a higher mean Story Recall compound score, caused by a higher score both on Store Recall Immediate and Story Recall Delayed.

Explorative GLMs were performed to investigate whether training A differed in outcome from training B. On none of the compound scores from pre- to post-testing and from post- to follow-up testing were significant differences found between the two training groups.

In sum, looking at the immediate training effects (pre- to post-testing), group A showed an increase in verbal STM compared with the control group. This gain was maintained at least till the follow-up session. Between post- and follow-up testing, groups A and B showed an increase compared with the control group on compound scores of visual STM, Scholastic abilities and Story Recall, while group B also improved on visuo-spatial WM. Looking at the effects from pre-intervention to follow-up, group A and group B obtained both higher scores on compound scores of Scholastic abilities and on Story Recall compared with the control group, while group A additionally obtained a higher score on visuo-spatial WM.

Table 1a Mean scores (SD) for primary outcome measures of STM and WM performance in the pre-, post- and follow-up condition for group A, B and the control group

	DR	NR	BlR	VP	BR	LR	SpS
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)
Group A							
Pre ($n = 41$)	20.34 (4.44)	13.91 (3.09)	27.37 (3.88)	16.74 (3.67)	11.67 (3.89)	12.24 (3.01)	17.18 (5.52)
Post ($n = 41$)	21.60 (4.86)	15.21 (3.40)	27.30 (3.86)	19.26 (4.35)	11.62 (4.31)	14.16 (3.69)	19.72 (6.27)
Follow up ($n = 39$)	21.38 (5.52)	15.46 (4.12)	28.21 (4.53)	19.95 (4.82)	12.00 (5.21)	14.77 (4.53)	20.15 (6.26)
Group B							
Pre ($n = 27$)	20.37 (5.31)	14.61 (3.40)	26.15 (4.80)	15.65 (4.01)	12.17 (4.43)	12.20 (3.47)	16.78 (6.19)
Post ($n = 26$)	21.11 (4.86)	15.52 (3.50)	25.05 (4.31)	18.59 (4.12)	12.74 (4.72)	13.98 (3.17)	18.46 (5.32)
Follow up ($n = 25$)	20.24 (4.65)	15.60 (3.46)	26.52 (4.77)	19.24 (4.28)	12.24 (5.07)	14.52 (4.78)	19.56 (5.54)
Control group							
Pre ($n = 27$)	20.22 (3.39)	13.91 (1.95)	25.91 (4.47)	15.61 (3.45)	12.33 (3.83)	12.15 (2.68)	16.96 (3.66)
Post ($n = 26$)	20.44 (3.59)	14.52 (2.14)	26.46 (4.41)	18.56 (3.64)	11.42 (3.89)	13.81 (2.65)	18.69 (5.27)
Follow up ($n = 25$)	20.32 (4.32)	14.84 (3.64)	25.48 (4.22)	19.12 (5.38)	11.88 (3.64)	14.52 (2.80)	17.92 (5.18)

BlR, Block Recall; BR, Backward Digit Recall; DR, Digit Recall; LR, Listening Recall; NR, Nonword Recall; SpS, Spatial Span; STM, short-term memory; VP, Visual Patterns test; WM, working memory.

Table 1b Mean scores (SD) for secondary outcome measures in the pre-, post- and follow-up condition for group A, B and the control group

	Ar	Re	SRi	SRd	St	R
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*M* (SD)
Group A						
Pre ($n = 41$)	71.19 (25.64)	65.51 (17.62)	11.00 (3.40)	8.95 (3.53)	46.05 (19.86)	35.37 (6.26)
Post ($n = 41$)	73.94 (26.19)	72.23 (18.14)	14.11 (4.31)	12.66 (4.36)	36.22 (16.55)	36.24 (5.99)
Follow up ($n = 39$)	76.56 (25.24)	75.31 (18.74)	15.45 (5.04)	13.44 (5.19)	34.76 (18.40)	35.31 (7.30)
Group B						
Pre ($n = 27$)	82.55 (25.46)	74.09 (19.23)	11.22 (4.58)	9.78 (4.43)	45.89 (15.32)	32.70 (7.91)
Post ($n = 26$)	86.70 (25.54)	79.78 (19.55)	14.63 (5.32)	13.41 (5.76)	38.26 (14.61)	33.35 (7.57)
Follow up ($n = 25$)	89.84 (28.89)	83.16 (20.76)	16.24 (5.92)	15.16 (6.39)	32.98 (12.87)	33.16 (8.13)
Control group						
Pre ($n = 27$)	73.57 (25.22)	68.17 (18.62)	11.09 (3.60)	9.39 (3.15)	45.70 (16.48)	33.37 (5.29)
Post ($n = 26$)	76.40 (23.98)	73.26 (19.0)	14.15 (4.03)	13.06 (4.10)	39.60 (17.47)	35.58 (6.99)
Follow up ($n = 25$)	76.52 (25.48)	74.92 (19.43)	13.24 (4.03)	12.08 (3.82)	36.30 (20.07)	34.64 (7.73)

Ar, Arithmetic; Re, Reading; SRd, Story Recall delayed; SRi, Story Recall immediate; St, Stroop; R, Raven SPM.

Discussion

This study provides the first demonstration that WM can be effectively trained in adolescents with M-BID, a group known for its low WM capacity (e.g. Henry 2001; Van der Molen *et al.* 2007, 2009). Verbal STM improved significantly from pre- to post-testing in the group who received the adaptive training compared with the control group. The beneficial effect on verbal STM was maintained at follow-up and other effects became clear at that time as well. Both the adaptive and non-adaptive WM training led to higher scores at follow-up than at post-intervention on visual STM, arithmetic and

Table 2 General linear model with contrasts between group A and group B versus the control group for post-testing with pre-intervention score as covariate and for follow-up testing with post-intervention score as covariate with η^2_p, B and one-sided *P*. Significant effects in bold.

	Pre to post			Post to follow-up		
Compound measures	η^2_p	*B*	*P*	η^2_p	*B*	*P*
Verbal STM						
Training A – Control	**0.04**	**0.17**	**0.03**	0.02	0.16	0.12
Training B – Control	0.00	0.06	0.28	0.00	–0.05	0.62
Visual STM						
Training A – Control	0.00	–0.02	0.58	**0.03**	**0.19**	**0.05**
Training B – Control	0.01	–0.12	0.84	**0.04**	**0.24**	**0.03**
Verbal WM						
Training A – Control	0.00	0.09	0.24	0.00	–0.05	0.62
Training B – Control	0.01	0.15	0.14	0.01	–0.12	0.76
VS WM						
Training A – Control	0.00	0.81	0.25	0.02	1.57	0.08
Training B – Control	0.00	–0.2	0.56	**0.05**	**2.44**	**0.02**
Schol. ab.						
Training A – Control	0.00	0.01	0.45	**0.04**	**0.14**	**0.03**
Training B – Control	0.00	0.01	0.43	**0.05**	**0.16**	**0.02**
Story Recall						
Training A – Control	0.00	–0.05	0.58	**0.05**	**0.56**	**0.02**
Training B – Control	0.00	–0.02	0.50	**0.09**	**0.85**	**0.04**
Resp. inhib.						
Training A – Control	0.01	–3.48	0.12	0.01	3.22	0.81
Training B – Control	0.00	–1.34	0.34	0.00	–0.43	0.46
Fluid int.						
Training A – Control	0.00	–0.41	0.62	0.00	–0.71	0.30
Training B – Control	0.01	–1.57	0.86	0.01	–1.17	0.23

The lower the score on Response inhibition, the better the performance.
η^2_p, Partial Eta squared: 0.01 = small effect; 0.06 = medium effect, 0.14 = large effect (Cohen, 1988); *B*, non-standardised regression coefficient; Fluid int., Fluid intelligence; Resp. inhib., Response inhibition; Schol. ab, Scholastic abilities; STM, short-term memory; VS WM, Visuo-spatial WM; WM, working memory.

story recall compared with the control condition. In addition, the non-adaptive training group showed a significant increase in visuo-spatial WM capacity. These effects were evaluated further by performing pre-test versus follow-up test comparisons. These comparisons indicated that arithmetic and the ability to recall stories improved significantly from pre- to follow-up testing both for the adaptive and non-adaptive training groups but not for the control group. In addition, when comparing between pre- and follow-up testing, it was the adaptive training group that gained in visuo-spatial WM capacity. Apparently, the WM training 'Odd Yellow' succeeded in increasing the ability to both simultaneously process and store information at the short term, leading to small, but significant effects on other, scholastic, and everyday tasks like arithmetic and recalling a story. This is in line with previous training studies (e.g. Klingberg *et al.* 2005; Holmes *et al.* 2009). Furthermore, it means that individuals with mild to borderline ID are able to generalise what they have learnt, in contrast to what is suggested (e.g. Park & Gaylord-Ross 1989; Bebko & Luhaorg 1998).

Increased effects at follow-up compared with immediate effects following WM training have been reported before (Klingberg *et al.* 2005; Van 't Hooft *et al.* 2007; Holmes *et al.* 2009). For example, Holmes *et al.* (2009) found a similar late effect on an arithmetic test in typically developing children with low WM. The authors argued that such a late effect is unsurprising 'as any improved cognitive support for learning caused by training would be expected to take some time to work its way through

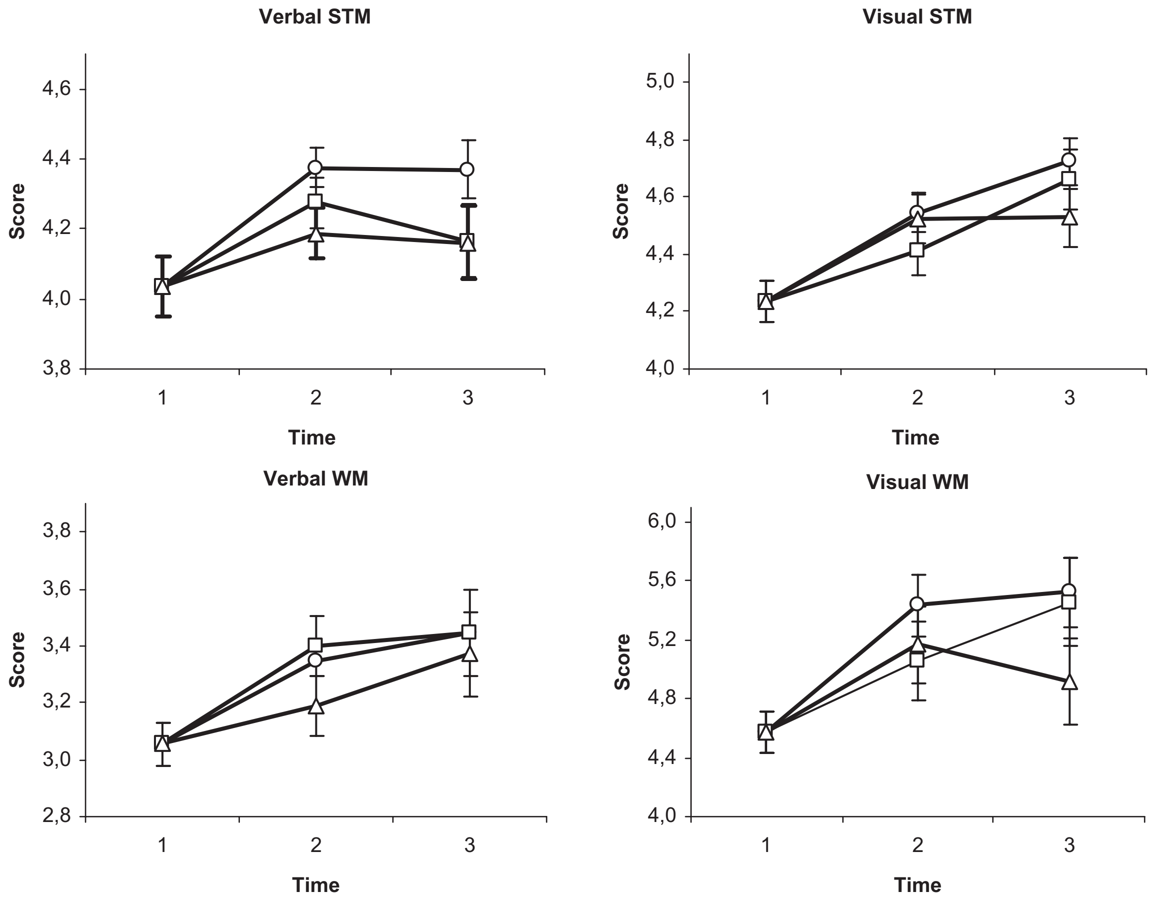


Figure 4 Adjusted mean with standard error at pre-testing (time 1), post-intervention (time 2) and at follow-up (time 3) corrected for differences in baseline score, with the baseline score set to the common average for all groups. STM, short-term memory; WM, working memory.

to significant advances in performance on standardised ability tests' (p. F13).

The visuo-spatial WM training in this study showed positive effects on a visuo-spatial WM task that was not trained and on other cognitive activities like arithmetic; it had no effect on verbal WM. This contrasts the results of Thorell *et al.* (2009) who did find transfer effect of their visuo-spatial WM training to verbal WM outcome measures. Their training consisted of 23 sessions of 15 min each, while our training consisted of 15 sessions of 6 min each. It might be the difference in intensity which led to the different outcomes. It is worthwhile exploring if verbal WM can be improved in adolescents with M-BID, as especially verbal WM is weak in this population (Maehler & Schuchardt 2009; Van der Molen *et al.* 2009). This WM training should then possibly comprise both visual and verbal WM as the visuo-spatial WM training exerted its beneficial effect on an important target domain, i.e. scholastic abilities.

In the current study, we did not obtain training effects on fluid intelligence and response inhibition. Holmes *et al.* (2009) failed to observe effects on fluid intelligence as well (this study did not include a response inhibition task). Studies have shown that WM processes influence fluid intelligence in both typically developing adults (e.g. Engle RW *et al.* 1999) and children (e.g. Swanson 2008), as well as inhibition in typically developing adults (e.g. Friedman *et al.* 2006) and in children (e.g. Swanson 2008). Furthermore, Klingberg *et al.* (2005)

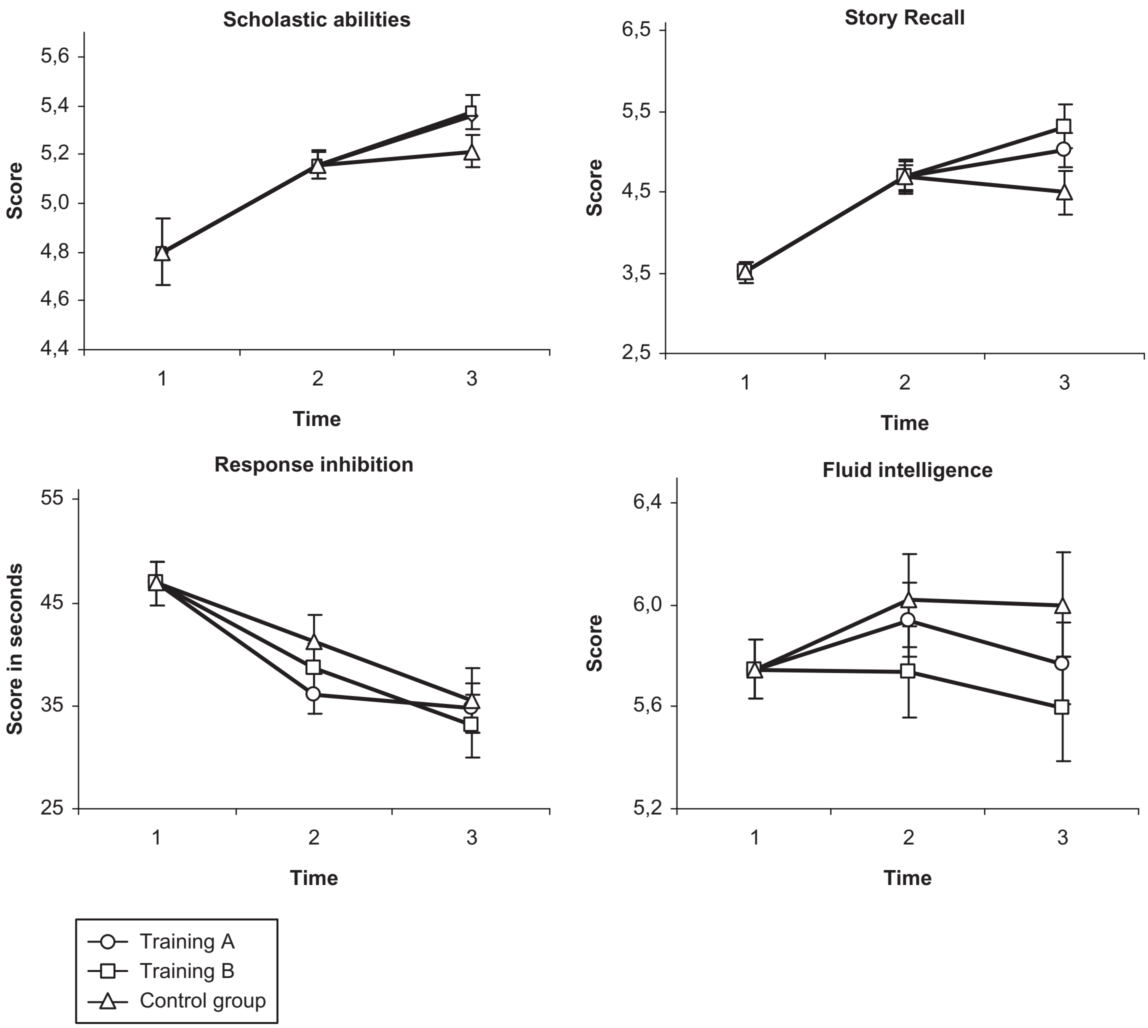


Figure 4 Continued.

observed effects on both fluid intelligence and response inhibition after their WM training in children with ADHD, but only at post-testing and not at follow-up. Jaeggi *et al.* (2009) reported a significant improvement in fluid intelligence in WM-trained university students as well (there was no follow-up measurement), although questions are being raised about the way the authors measured fluid intelligence (Moody 2009). It is hard to explain the difference in outcome. For example, the amount of training sessions and the duration of each session were comparable in the study of Holmes *et al.* and Klingberg *et al.* (in both studies 25 sessions of 30 min each), and Jaeggi *et al.* (19 sessions or less of 25 min each). Populations were all quite different: university students (Jaeggi *et al.*), children with ADHD (Klingberg *et al.*), typically developing children with low WM capacity (Holmes *et al.*) and adolescents with M-BID (this study). Clearly, more training studies are needed, varying in populations and training intensity to see what the exact relationship between WM, fluid intelligence and inhibition is for specific groups and if these functions can be effectively trained.

As Fig. 3 shows, children in the adaptive training group (A) obtained span scores above 3.5 during most sessions. Surprisingly however, the effects of the adaptive training (A) and the non-adaptive training (B, with a maximum span possibility of 2) were basically similar. This contrasts with previous studies showing larger effects of adaptive relative to non-adaptive training (Klingberg *et al.* 2005; Holmes *et al.* 2009). The population in those studies consisted of typically developing children either with low WM capacity (Holmes *et al.*), or with ADHD (Klingberg *et al.*). The fact that we did not find a difference between the two training versions might either be caused by the lower intelligence of the children, which possibly makes a low intensity training an effective one, or by the low training intensity preventing to exhibit a difference between the two versions.

Compared with the aforementioned training studies, frequency and intensity in our training were relatively low. Nevertheless, we did find improvements on several outcome measures. Therefore, it seems that our training 'Odd Yellow' was adequate for adolescents with M-BID. However, the effect sizes were mostly small and we failed to observe an effect on verbal WM. Most likely, this is due to the low intensity of the training. Therefore, it would be worthwhile exploring the possibilities to intensify the training for this population with longer and more sessions. Special attention should then be given to the attractiveness of the training which could for example be arranged by packing it into a game (Prins *et al.* in press). However, the results of the current study are encouraging in that apparently WM, a central and important cognitive aspect, can be trained effectively with a fanning out effect on scholastic and other everyday tasks in a cognitive weak and therefore vulnerable group of people; children with mild to borderline ID.

Acknowledgments

This research was supported by grants from Stichting Steunfonds 's Heeren Loo, Stichting tot Steun VCVGZ and 's Heeren Loo Zorggroep, The Netherlands. The authors would like to thank all children who took part in the research, and staff at the following schools: Arkelstein, De Noordhoek, Emaus College, Joannesschool Bemmel, Praktijkschool Woerden, Praktijkschool Helmond, Thomas A Kempis College. Furthermore thanks to Judith Arendsen, Marieke Blok, Marjet Brink, Marion Derksen, Lindsay Lacet, Marjolein Sterk and Ellen Stevens for assistance with data collection.

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. (2007) *Automated Working Memory Assessment*. Harcourt Assessment, London.

Alloway T. P., Gathercole S. E., Kirkwood H. & Elliott J. (2009) The cognitive and behavioral characteristics of children with low working memory. *Child Development* **80,** 606–21.

American Psychiatric Association (2000) *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR*. American Psychiatric Association., Washington, DC.

Baddeley A. (1986) *Working Memory*. Clarendon Press, Oxford.

Bebko J. M. & Luhaorg H. (1998) The development of strategy use and metacognitive processing in mental retardation: some sources of difficulty. In: *Handbook of Mental Retardation and Development* (eds J. A. Burack, R. M. Hodapp & E. Zigler), pp. 382–407. Cambridge University Press, Cambridge.

Brus B. T. & Voeten M. J. M. (1973) *Eén minuut test [One minute reading test]*. Berkhout, Nijmegen.

Cohen J. (1988) Statistical power analysis for the behavioural sciences (2nd ed.), Lawrence Erlbaum Associates: New Jersey.

Comblain A. (1994) Working memory in Down syndrome: training the rehearsal strategy. *Down's Syndrome, Research and Practice* **2,** 123–6.

Conners F. A., Rosenquist C. J. & Taylor L. A. (2001) Memory training for children with Down syndrome. *Down's Syndrome, Research and Practice* **7,** 25–33.

Conners F. A., Rosenquist C. J., Arnett L., Moore M. S. & Hume L. E. (2008) Improving memory span in children with Down syndrome. *Journal of Intellectual Disability Research* **52,** 244–55.

Cowan N. (1999) An embedded-processes model of working memory. In: *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (eds A. Miyake & P. Shah), pp. 62–101. Cambridge University Press, Cambridge.

Craik F. I. M., Woncur G., Palmer H., Binns M. A., Edwards M., Bridges K. *et al.* (2007) Cognitive rehabilitation in the elderly: effects on memory. *Journal of the International Neuropsychological Society* **13,** 132–42.

De Vos T. (1992) *Tempo Test Rekenen [Tempo Test Arithmetic]*. Swets Test Publishers, Lisse.

Della Sala S., Gray C., Baddeley A. & Wilson L. (1997) *Visual Patterns Test: A New Test of Short-Term Visual Recall*. Thames Valley Test Company, Suffolk.

Engle R. E., 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 functions of the prefrontal cortex. In: *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (eds A. Miyake & P. Shah), pp. 102–34. Cambridge University Press, Cambridge.

Engle R. W. & Nagle R. J. (1979) Strategy training and semantic encoding in mildly retarded children. *Intelligence* **3,** 17–30.

Engle R. W., Tuholski S. W., Laughlin J. E. & Conway A. R. A. (1999) Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology. General* **128,** 309–31.

Friedman N. P., Miyake A., Corley R. P., Young S. E., DeFries J. C. & Hewitt J. K. (2006) Not all executive functions are related to intelligence. *Psychological Science* **17,** 172–9.

Garon N., Bryson S. E. & Smith I. M. (2008) Executive functioning in preschoolers: a review using an integrative framework. *Psychological Bulletin* **134,** 31–60.

Gathercole S. E. & Pickering S. J. (2000) Assessment of working memory in six- and seven-year-old children. *Journal of Educational Psychology* **92,** 377–90.

Hammes J. G. W. (1978) *De Stroop Kleur-Woord Test [The Stroop Color-Word Test]*. Swets & Zeitlinger, Lisse.

Henry L. A. (2001) How does the severity of a learning disability affect working memory performance? *Memory* **9,** 233–47.

Henry L. A. & Miller S. (1993) Why does memory span improves with age? A review of the evidence for the two current hypotheses. *European Journal on Cognitive Psychology* **5,** 241–87.

Hitch G. J., Towse J. N. & Hutton U. (2001) What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology. General* **130,** 184–98.

Holmes J., Gathercole S. E. & Dunning D. L. (2009) Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science* **12,** F9–15.

Jaeggi S. M., Buschkuehl M., Jonides J. & Perrig W. J. (2009) Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 6829–33.

Kellas G., Ashcraft M. H. & Johnson N. S. (1973) Rehearsal processes in the short-term memory performance of mildly retarded adolescents. *American Journal of Mental Deficiency* **77,** 670–9.

Klingberg T., Forssberg H. & Westerberg H. (2002) Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology* **24,** 781–91.

Klingberg T., Fernell E., Olesen P. J., Johnson M., Gustafsson P., Dahlström 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,** 177–86.

Kurtz M. M., Seltzer J. C., Shagan D. S., Thime W. R. & Wexler B. E. (2007) Computer-assisted cognitive remediation in schizophrenia: what is the active ingredient? *Schizophrenia Research* **89,** 251–60.

Laws G., MacDonald J., Buckley S. & Broadley I. (1995) Long-term maintenance of memory skills taught to children with Down's Syndrome. *Down's Syndrome, Research and Practice* **3,** 103–9.

Laws G., MacDonald J. & Buckley S. (1996) The effects of a short training in the use of a rehearsal strategy on memory for words and pictures in children with Down syndrome. *Down's Syndrome, Research and Practice* **4,** 70–8.

Lezak M. D. (1995) Neuropsychological Assessment (3rd ed.), New York, NY: Oxford University Press.

Maehler C. & Schuchardt K. (2009) Working memory functioning in children with learning disabilities: does intelligence make a difference? *Journal of Intellectual Disability Research* **53,** 3–10.

Minear M. & Shah P. (2006) Sources of working memory deficits in children and possibilities for remediation. In: *Working Memory and Education* (ed. S. Pickering), pp. 273–307. Academic Press, London.

Moody D. E. (2009) Can intelligence be increased by training on a task of working memory? *Intelligence* **37,** 327–8.

Park H. S. & Gaylord-Ross R. (1989) A problem-solving approach to social skills training in employment settings with mentally retarded youth. *Journal of Applied Behaviour Analysis* **22,** 373–80.

Pickering S. J. & Gathercole S. E. (2001) *Working Memory Test Battery for Children*. Psychological Corporation, London.

Prins P. J. M., Dovis S., Ponsioen A., Ten Brink E. & Van der Oord S. (in press) Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? *Cyberpsychology and Behavior*.

Raven J. C., Court J. H. & Raven J. (1996) *Manual for Raven's Standard Progressive Matrices and Vocabulary Scales*. Oxford Psychologists Press, Oxford.

Sabornie E. J., Cullinan D., Osborne S. S. & Brock L. B. (2005) Intellectual, academic, and behavioral functioning students with high-incidence disabilities: a cross-categorical meta-analysis. *Exceptional Children* **72,** 47–63.

Simonoff E., Pickles A., Chadwich O., Gringras P., Wood N., Higgins S. *et al.* (2006) The Croydon assessment of learning study: prevalence and educational identification of mild mental retardation. *Journal of Child Psychology and Psychiatry* **47,** 828–39.

Swanson H. L. (2008) Working memory and intelligence in children: what develops? *Journal of Educational Psychology* **100,** 581–602.

Thorell L. B., Lindqvist S., Berman Nutley S., Bohlin G. & Klingberg T. (2009) Training and transfer effects of executive functions in preschool children. *Developmental Science* **12,** 106–13.

Van der Molen M. J. (2007) *Een Kort Verhaaltje Om Te Onthouden [A Short Story to Remember]*. Utrecht University (internal publication), Utrecht.

Van der Molen M. J., Van Luit J. E. H., Jongmans M. J. & Van der Molen M. W. (2007) Verbal working memory in children with mild intellectual disabilities. *Journal of Intellectual Disability Research* **51,** 162–9.

Van der Molen M. J., Van Luit J. E. H., Jongmans M. J. & Van der Molen M. W. (2009) Memory profiles in children with mild intellectual disabilities: strengths and weaknesses. *Research in Developmental Disabilities* **30,** 1237–47.

Van der Molen M. J., Van Luit J. E. H., Jongmans M. J. & Van der Molen M. W. (in press) Everyday memory and working memory in children with mild intellectual disabilities. *American Journal on Intellectual and Developmental Disabilities* (in press).

Van 't Hooft I., Andersson K., Bergman B., Sejersen T., Von Wendt L. & Bartfai A. (2007) Sustained favorable effects of cognitive training in children with acquired brain injuries. *NeuroRehabilitation* **22,** 109–16.

Accepted 22 March 2010