Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of Experimental Child Psychology 96 (2007) 249–283
Journal of Experimental Child Psychology
www.elsevier.com/locate/jecp

The influence of working memory on reading growth in subgroups of children with reading disabilities

H. Lee Swanson *, Olga Jerman

Department of Educational Psychology, University of California, Riverside, Riverside, CA 92521, USA

Received 10 April 2006; revised 6 December 2006

Abstract

This 3-year longitudinal study determined whether (a) subgroups of children with reading disabilities (RD) (children with RD only, children with both reading and arithmetic deficits, and low verbal IQ readers) and skilled readers varied in working memory (WM) and short-term memory (STM) growth and (b) whether growth in an executive system and/or a phonological storage system mediated growth in reading performance. A battery of memory and reading measures was administered to 84 children (11–17 years of age) across three testing waves spaced 1 year apart. The results showed that skilled readers yielded higher WM growth estimates than did the RD groups. No significant differentiation among subgroups of children with RD on growth measures emerged. Hierarchical linear modeling showed that WM (controlled attention), rather than STM (phonological loop), was related to growth in reading comprehension and reading fluency. The results support the notion that deficient growth in the executive component of WM underlies RD.

Keywords: Working memory; Learning disabilities; Reading; Fluency; Processing; Executive processing; Phonological processing

Introduction

Cognitive impairments are important correlates of functional outcome in children with reading disabilities (RD). Of these impairments, working memory (WM) has been the

* Corresponding author.
E-mail address: lee.swanson@ucr.edu (H.L. Swanson).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2006.12.004

focus of extensive research efforts because it plays a central role in several domains of cognition, including language comprehension, fluid intelligence, writing, and arithmetic (Gathercole, Alloway, Willis, & Adams, 2006; for a review, see also ; Swanson & Siegel, 2001). Furthermore, WM impairments have been related to specific aspects of RD such as problems in reading comprehension (Swanson, 1999). Thus, there is evidence that WM impairments may play a critical role in mediating some of the academic problems in children with RD (e.g., de Jong, 1998; Gathercole et al., 2006; Swanson & Berninger, 1995; Willcutt, Pennington, Olson, Chhabildas, & Hulslander, 2005). Although WM is integrally related to a number of academic behaviors, relatively few studies have been undertaken to systematically explore whether growth in WM underlies growth in reading performance.

One possible framework to capture growth in WM and its influence on reading is Baddeley's (1986, 1996, 2000) multicomponent model. Baddeley and Logie (1999) described WM as a limited capacity central executive system that interacts with a set of two passive store systems used for temporary storage of different classes of information: the speech-based phonological loop and the visual sketchpad. The phonological loop is responsible for the temporary storage of verbal information. Within the phonological store, items are held for a limited duration and are maintained via the process of articulation. The visual sketchpad is responsible for the storage of visual–spatial information over brief periods and plays a key role in the generation and manipulation of mental images. Both storage systems are in direct contact with the central executive system. The central executive system is considered to be primarily responsible for coordinating activity within the cognitive system, but it also devotes some of its resources to increase the amount of information that can be held in the two subsystems (Baddeley & Logie, 1999). A recent formulation of the model (Baddeley, 2000) also includes a temporary multimodal storage component called the episodic buffer.

In one of the few developmental studies testing Baddeley's WM model in children with RD, Swanson (2003) compared participants with RD with skilled readers ($N = 226$) across four age groups (7, 10, 13, and 20 years) WM tasks. As expected, participants with RD were inferior to skilled readers at all age levels, but these differences increased with age. The age-related differences between ability groups were maintained under cuing conditions and when reading and arithmetic skills were partialed from the analysis. Skilled readers showed age-related increases in WM, whereas the trajectory of growth for children with RD showed minimal age-related changes in span level across ages 7–20 years.

These findings raise the question as to whether constraints in WM growth constrain growth in reading. Although the literature is clear that impairments in WM in children with RD are related to reading performance, whether growth in WM underlies poor growth in reading skills has not been tested empirically. Gathercole, Tiffany, Briscoe, and Thorn (2005) conducted one of the closest studies on this issue. Although their research was not focused on RD per se, they initiated a longitudinal study that investigated the cognitive skills and scholastic attainments of 8-year-olds who were selected on the basis of deficits in the phonological loop at 4 years of age. The phonological loop was assessed by short-term memory (STM) tasks that tapped the recall of verbal information. The authors investigated whether deficits in phonological STM performance during this developmental period had a direct impact on children's attainment in the areas of language, arithmetic, and literacy. They reasoned that if phonological memory was a significant deterrent to learning, children with poor phonological memory should have low overall achievement. In fact, children who scored low on phonological STM measures assessed at 4 and 8 years of age performed at appropriate levels in all areas of vocabulary, language, number skills, and literacy. Gathercole and colleagues concluded that children did not experience learning difficulties in key domains over the early school years that could be attributed to poor phonological memory. However, they found that the relation with WM and literacy was significant ($r = 56$).

As an extension of this earlier work, our aim was to determine whether poor phonological memory and/or executive functioning may underlie subsequent growth in literacy in older children. It is possible that growth on academic tasks is more sensitive to phonological memory constraints in older children as language abilities approach adult levels. There is also some literature suggesting that tasks related to the phonological loop (i.e., STM) and the executive system are not clearly distinguishable in young children (i.e., high intercorrelation between WM and STM measures [Alloway, Gathercole, Willis, & Adams, 2004; Gathercole, Pickering, Ambridge, & Wearing, 2004]); therefore, the phonological system may play a more important role in later academic performance when older children with RD are included in the sample. Furthermore, Engle, Tuholski, Laughlin, and Conway (1999) suggested that STM and WM might not be distinguishable in young children whose rehearsal processes are unstable (i.e., used inconsistently). For example, several studies show weak to moderate correlations between memory span and articulation rates in children 8 years of age or younger (for a review, see Gathercole, 1998, pp. 6–7). This occurs because younger children's time-based strategies (i.e., rehearsal) are unstable (Gathercole, 1998). Thus, we selected an age level where memory strategies are stable to better assess the processes that moderate the relation among WM, STM, and reading.

Assuming that WM does underlie reading growth, two models were tested in the current study. One is that poor phonological STM constrains growth on literacy measures.[1] Several studies that have compared skilled readers with children with RD assume that STM measures capture a subset of WM performance—the use and/or operation of the phonological loop (for a comprehensive review, see Swanson & Siegel, 2001). This is because successful performance on STM measures draws on two major components of the phonological loop: a speech-based phonological input store and a rehearsal process (Baddeley, 1986). Research to date on STM indicates that children with RD rehearse less and perform more poorly than do skilled readers on tasks requiring the short-term retention of order information (e.g., O'Shaughnessy & Swanson, 1998), suggesting inefficient

[1] The distinctions made between the central executive system and a passive storage system (i.e., the phonological loop) in Baddeley and Logie's (1999) model in some ways parallel the distinctions made between WM and STM. WM is referred to as a processing resource of limited capacity involved in the preservation of information while processing the same or other information (e.g., Baddeley & Logie, 1999; Engle et al., 1999; Just, Carpenter, & Keller, 1996; Oberauer, Sub, Wilhelm, & Wittman, 2003) In contrast, STM typically involves situations where small amounts of material are held passively (i.e., minimal resources from long-term memory are activated to interpret the tasks such as digit or word span tasks) and then are reproduced in a sequential fashion. That is, participants are asked only to reproduce the sequence of items in the order they were presented (e.g., Daneman & Carpenter, 1980; Dempster, 1985; Engle et al., 1999). Although STM and WM share a close relation (i.e., transient memory [eg., Engle et al., 1999, reported a correlation of .70]), WM tasks are assumed to place heavy demands on an executive system and, therefore, to tap mental resources not relied on when performing STM tasks (e.g., Daneman & Carpenter, 1980; Engle et al., 1999). In contrast, the phonological loop is associated with STM because it involves two major components discussed in the STM literature: a speech-based phonological input store and a rehearsal process (for reviews, see Baddeley, 1986; Gathercole, 1998).

use of the phonological rehearsal process (cf. Henry & Millar, 1993). Developmental differences in the phonological loop, therefore, might be expected to influence some aspects of growth in literacy (e.g., Leather & Henry, 1994).

Given that the phonological loop is partly controlled by the central executive system (i.e., the executive system shares some variance with the phonological loop), the development of reading may also be directly related to the controlling functions of the central executive system itself. Thus, in contrast to the previous model, the second model views age-related changes in the executive attention component of WM as underlying children's reading performance. However, the model assumes that there is variance unique to particular subsystems of WM (executive processing and phonological loop) and that both subsystems contribute to children's reading. Thus, in contrast to the aforementioned hypothesis suggesting that phonological processes play the dominant role in the mediating effects of WM on children's reading, the second model assumes that executive processes also play a major role in mediating reading performance. There are clear expectations for this second model. The influence of WM on measures of reading follows automatically with improvements in controlled attention. Specifically, if growth in reading is explained by activities related to controlled attention, growth in WM should be correlated with growth in reading.

Subgroup profiles

Because STM and WM tasks are strongly correlated, it may be difficult to determine whether reading deficits and poor WM performance are due primarily to a system related to storage (i.e., STM) or to a system that taps both storage and executive processing (i.e., WM). One means to untangle these possible sources of difficulties in reading is to compare subgroups of children with RD. Evidence for the contribution of different memory systems to reading deficits is suggested when different types of reading problems exhibit different profiles. Some studies have suggested that children who have combined deficits in reading and arithmetic (referred to as the comorbid group in the current study) reflect more generalized deficits related to the executive system than do children with RD only, that is, children who have deficits related only to reading (e.g., van der Sluis, van der Leij, & de Jong, 2005). Van der Sluis and colleagues (2005) suggested that children with RD do not exhibit problems in executive function, whereas children who show reading plus arithmetic difficulties exhibit problems in executive processing. Thus, of particular interest in the current study was to examine whether performance of children identified with only RD reflects deficits in the development of phonological STM, whereas performance of children with comorbid deficits in both reading and arithmetic reflects developmental deficits in the executive system.

The rationale behind these hypotheses is as follows. There is tremendous consensus in the literature that poor phonological coding is related to poor word recognition (e.g., Shankweiler et al., 1995; for reviews, see also Stanovich, 1990; Stanovich & Siegel, 1994); thus, a significant relation between STM and word recognition would be expected in the current study. Furthermore, the relation between WM and reading is well established (Daneman & Merikle, 1996; Engle, Cantor, & Carullo, 1992). What is unclear from the literature, however, is whether specific deficits in the phonological storage and/or controlled attention underlie differences between low reading groups on WM measures. It seemed to us that if the WM system plays its major role in more generalized academic problems (i.e., below average word recognition and arithmetic), children with comorbid deficits would experience memory problems related to the executive system. In contrast, children with specific deficits in reading only would be impaired on phonological STM (storage) tasks.

Purpose

The current study was designed to test whether growth in WM underlies growth in reading measures in subgroups of children with RD. The study addressed two questions:

1. Do different growth patterns in WM emerge in subgroups of children with RD and skilled readers?
2. Does an executive system (controlled attention) and/or a phonological storage system account for growth in reading performance?

Based on cross-sectional studies, we predicted that growth in WM underlies reading growth in children with RD only and children with more general academic deficits. We tested whether growth is stable (no change) or linear across age. We also tested whether more generalized WM deficits (controlled attention) and storage deficits underlie memory growth in the comorbid group when compared with children with RD only. Children with RD only and children with comorbid deficits were also compared with children who had comparable reading deficits but were low in verbal intelligence. Because more generalized deficits in achievement (problems in both reading and arithmetic) may be associated with verbal IQ, it was necessary to determine whether differences in WM growth are merely an artifact of verbal ability. Thus, also included in the comparison were children with low reading and low verbal IQ scores (referred to in this study as low verbal IQ readers). Nation, Adams, Bowyer-Crane, and Snowling (1999) suggested that the relation between WM and reading (i.e., comprehension) is explained by verbal intelligence and not by executive processing skills. They argued that low verbal IQ readers have a weakness in verbal skills that restricts their ability to store verbal information in STM. Likewise, Stothard and Hulme (1992) suggested that WM differences between good and poor readers (comprehenders) would be eliminated if differences in verbal IQ were controlled; however, no evidence was provided to support this prediction.

Method

Participants

Participants in this study were 84 children selected from a school in Southern California. The mean age of the sample in the first testing wave was 11.99 years ($SD = 2.57$), and the sample was composed of 28 girls and 56 boys. Of these participants, 66 were Anglo-American, 8 were African American, 5 were Hispanic, and 5 were of mixed heritage. All children were from middle-upper class to upper class homes. Fully 80% of the participants' fathers had a bachelor's degree, and 70% of their mothers had a bachelor's degree.

In addition, 30% of the fathers and 10% of the mothers in the sample had a graduate degree (e.g., M.A., Ph.D., M.D.).

The criterion for the average reading group was a score above the 45th percentile in reading word recognition as well as an IQ score greater than 90. Classification of subgroups of children with RD was based on standard scores from the word reading and arithmetic subtests of the Wide Range Achievement Test-3 (WRAT-3) (Wilkinson, 1993). Operational criteria for children defined as the RD-only group were based on (a) IQ scores (in this case scores on the Raven Progressive Matrices test and the verbal section of the Wechsler Intelligence Scales for Children–Third Edition [WISC-III]) greater than 90, (b) reading scores at or below the 25th percentile, and (c) arithmetic calculation scores above the 25th percentile. Children with average fluid intelligence (scores on the Raven Colored Progressive Matrices test [Raven, 1976]) and verbal IQs, but whose performance in both word recognition and calculation were below the 25th percentile, were classified as children with RD with comorbid deficits. Children whose fluid intelligence was in the average range, but who tested low (<25th percentile) in both arithmetic and word recognition and obtained low verbal IQ scores (<85), were defined as low verbal IQ readers. All children were monolingual in English. These cutoff score criteria for children with RD only matched the operational definition of RD outlined by Siegel and Ryan (1989).

Therefore, participants were divided into four ability groups: children with RD only (low word recognition and average arithmetic, $n = 18$), children with both reading and calculation deficits (comorbid group, $n = 18$), skilled readers (high word recognition and high arithmetic, $n = 23$), and low verbal IQ readers (low verbal IQ, low word recognition, and low arithmetic, $n = 25$). Children in each reading group were matched as closely as possible on socioeconomic status (SES), ethnicity, and chronological age. Significant differences (ps $< .05$) emerged related to gender, $\chi^2(3, N = 84) = 10.47$, $p < .05$, ethnicity, $\chi^2(15, N = 84) = 27.67$, $p < .05$, and chronological age, $F(3, 77) = 3.22$, $p < .05$. We investigated whether gender and ethnicity influenced the results, and it was found to be nonsignificant (ps $> .05$). Thus, the results were pooled across gender and ethnicity. However, chronological age was used as a covariate when comparing the four ability groups on growth.

Subgroup comparisons on measures of word recognition, $F(3, 77) = 34.43$, $p < .001$, and arithmetic, $F(3, 77) = 16.11$, $p < .0001$, were significant. A Tukey's test indicated that significant differences (ps < 05) emerged on standard scores of word recognition (i.e., WRAT-III: low verbal IQ readers = comorbid group = children with RD only < skilled readers) and arithmetic (i.e., WRAT-III: low verbal IQ readers = comorbid group < children with RD only = skilled readers). The reading groups differed significantly on verbal IQ measures, $F(3, 77) = 19.47$, $p < .0001$. A Tukey's test indicated that significant subgroup differences (ps $< .05$) emerged on verbal IQ scores (i.e., low verbal IQ readers < children with RD only = comorbid group < skilled readers). Significant differences emerged between the groups on the Raven Progressive Matrices test standard scores, $F(3, 77) = 2.88$, $p < .05$. A Tukey's test indicated that significant subgroup differences (ps $< .05$) emerged on Raven test scores (i.e., low verbal IQ readers = comorbid group < children with RD only = skilled readers). Thus, when comparing children with RD only with the comorbid group, Raven test scores were partialed from the analysis (see Table 4 later).

Tasks and materials

The battery of group and individually administered tasks is described here. Experimental tasks are described in more detail than are published and standardized tasks. All experimental tasks included two practice trials.

Classification measures

Reading

Word recognition was assessed by the reading subtest of the WRAT-3. The task provided a list of words of increasing difficulty. The child was required to read the words until five consecutive errors occurred. The dependent measure was the number of words read correctly.

Arithmetic computation

The arithmetic subtest from the WRAT-3 was administered. The dependent measure was the number of problems solved correctly, which yielded a standard score.

Verbal intelligence

Verbal IQ was derived from two subtests of the WISC-III: general information and vocabulary.

Fluid intelligence

To ensure that children did not suffer from general intellectual deficits, fluid intelligence was assessed by the Raven Colored Progressive Matrices test. Children were given a booklet with patterns displayed on each page, where each pattern revealed a missing piece. For each pattern, six possible replacement pattern pieces were displayed. Children were required to circle the replacement piece that best completed the patterns. After the introduction of the first matrix, children completed their booklets at their own pace. Patterns progressively increased in difficulty. The dependent measure (range = 0–36) was the number of problems solved correctly.

Short-term memory measures

Three measures of STM were administered: forward digit span, word span, and pseudoword span. The forward digit span subtest from the WISC-III was administered. This task required participants to recall and repeat in order sets of digits that were spoken by the experimenter. Each set increased in number. The word span and pseudoword span tasks were presented in the same manner as the forward digit span measure. The word span task was used previously by Swanson, Ashbaker, and Lee (1996). The word stimuli were one- or two-syllable high-frequency words. Children were read lists of common but unrelated nouns and then were asked to recall the words. Word lists gradually increased in set size, from a minimum of two words to a maximum of eight words. The pseudoword span task (Swanson & Berninger, 1995) used strings of nonsense words (one syllable long) that were presented one at a time in sets of two to six nonwords (e.g., zeb, vab; sme, pru, tri). The dependent measure for all STM measures was the largest set of items retrieved in the correct serial order (range = 0–7).

Working memory measures

We administered the backward digit span task from the WISC-III as a measure of WM. This measure has been used in previous studies as an index of WM (e.g., Gathercole et al., 2004). Also administered was an updating task. The third and fourth measures (digit/sentence span and rhyming) required children to hold increasingly complex information in memory while responding to a question about the task. The questions served as distracters to item recall because they reflected the recognition of targeted and closely related nontargeted items. A question was asked for each set of items, and the tasks were discontinued if the question was answered incorrectly or if all items within a set could not be remembered. The question required a simple recognition of new and old information and was analogous to the yes/no response feature of Daneman and Carpenter's (1980) task. For all WM and STM tasks, children were presented with two practice trials. After successful completion of the practice trials, children were provided with sets of items that increased in difficulty. Task descriptions follow.

Backward digit span

The backward digit span task from the WISC-III was administered. The task required participants to recall in reverse order sets of digits administered. There were two sets for each level. The dependent measure was the highest set of items recalled in the correct backward order (range = 0–14).

Updating

The experimental updating task was adapted from Morris and Jones (1990). A series of one-digit numbers in varying set lengths of nine, seven, five, and three digits were presented. No digit appeared twice in the same set. The experimenter told the child that the length may be long or short. The participant was then told that he or she should recall only the last three numbers presented. Each digit was presented at approximately 1-s intervals. After the last digit was presented, the participant was asked to name the last three digits in order. We assumed that to recall the last three digits in an unknown ($N = 3, 5, 7, 9$) series of digits, the order of old information must be kept available (previously presented digits) along with the order of newly presented digits. The dependent measure was the largest set in which three examples in the set were repeated correctly (range = 0–9).

Digit/Sentence span

The digit/sentence span task assessed the child's ability to remember numerical information embedded in a short sentence (Swanson, 1992, 1995). Before stimulus presentation, the child was shown a card depicting four strategies for encoding numerical information to be recalled. The pictures portrayed the strategies of rehearsal, chunking, association, and elaboration. The experimenter described each strategy to the child before administration of targeted items. After all strategies had been explained, the child was presented with numbers in a sentence context. For example, one item stated, "Now suppose somebody wanted to have you take them to the supermarket at 8 6 5 1 Elm Street." The numbers were presented at 2-s intervals followed by a process question: "What was the name of the street?" Then the child was asked to select a strategy from an array of four strategies that represented the best approximation of how he or she planned to practice the information for recall. Finally, the experimenter prompted the child to recall the numbers from the sentence in order. No further information about the strategies was provided. The child was allowed 30 s to remember the information. Recall difficulty for this task included sets of items that ranged from 2 digits (Set 1) to 14 digits (Set 9). The dependent measure was the largest set recalled correctly in which the process question was answered correctly (range = 0–9).

Rhyming

The purpose of the rhyming task was to assess the child's recall of acoustically similar words (Swanson, 1995). The child listened to sets of words that rhyme. Each successive word in the set was presented every 2 s. There were nine word sets that ranged from 2 to 14 monosyllabic words. The dependent measure was the number of sets recalled. Before recalling the words, the child was asked whether a particular word was included in the set. For example, the child was presented with the words "lip–slip–clip" and then was asked whether "ship or lip" was presented in the word set. If the child answered the question correctly, he or she was asked to recall the previously presented words (lip–slip–clip) in order. The dependent measure was the number of sets recalled correctly (range = 0–9).

Criterion measures

Although ability groups were categorized on measures of word recognition and arithmetic skills, it was important to follow growth on measures related to other forms of reading independent of the classification measures. We selected two areas: reading comprehension and reading fluency.

Reading comprehension

Reading comprehension was assessed by the passage comprehension subtest from the Woodcock Reading Mastery Test-Revised (WRMT-R) (Woodcock, 1998). The purpose of this task was to assess the child's comprehension of topic or subject meaning during reading activities. Comprehension is assessed using a cloze procedure that draws on the reading of short paragraphs. The dependent measure was the number of comprehension questions answered correctly.

Word fluency

Real word reading efficiency was assessed by the sight word efficiency subtest from the Test of Word Reading Efficiency (TOWRE) (Wagner & Torgesen, 1999). The subtest required oral reading of a list of 120 real words of increasing difficulty. The child was given 45 s to read aloud as many words as possible from a list of common words. The dependent measure was the number of words read correctly in 45 s.

Reliability

The reliability of the measures was computed for the total sample. Because the subsequent analysis relied on raw scores for analysis, the reliability of all raw score measures was calculated. As shown in Appendixes A, B, and C, Cronbach's alpha scores varied, with the majority of coefficients hovering around 70. The majority of reliability coefficients are within an acceptable range for basic research (~.70) (for a discussion, see Nunnally & Bernstein, 1994, pp. 264–265).

Statistical analysis

Repeated measures

Prior to modeling whether WM contributed to growth in reading comprehension or fluency, we used a mixed model, also referred to as hierarchical linear modeling (HLM) (Bryk & Raudenbush, 1992; Singer, 2002), with repeated measures to determine whether significant differences in level and growth emerged between ability groups. The method overcomes some of the limitations of traditional analysis of variance (ANOVA) repeated measures because it does not assume that an equal number of repeated observations are taken for each individual or that all individuals were measured at the same time point. The model also allowed us to use random effects to model the continuous functions of time. Furthermore, the HLM procedure does not require that missing data be ignored and provides a valid means to addressing standard errors (Littell, Henry, & Ammerman, 1998). In contrast to traditional ANOVA repeated measures, where significance is tested against the residual error, the fixed effects in mixed models are tested against the appropriate error terms as determined by the model specification. In addition, the HLM procedure has an advantage over traditional repeated measures ANOVA designs because the covariance structure can be modeled. In the current study, we found that the repeated measures model outlined by Littell and colleagues (1998), which used combined random effects between participants and an autoregressive covariance structure within participants, best captured the repeated measures model.

Growth

Another advantage of HLM analysis over more traditional methods (repeated measures ANOVA) for the study of growth is that it goes beyond group comparisons. Thus, we addressed whether WM was related to growth in reading in the total sample. To examine reading growth, we again used a random effects model (Singer, 2002). Age was the variable that represented the passage of time in our growth model. To interpret the results, we centered age at 13.85 years (sample mean chronological age at Wave 3) so that intercepts reflected the expected performance at that age. (It is important to note that slope remains the same whether the mean age is centered at Wave 1, Wave 2, or Wave 3; however, the correlation between the intercept and the slope does vary as a function of different centering.) Our growth model yielded parameter estimates that defined both the overall trajectory of the sample (fixed effects) and deviations in the overall trajectories (random effects). Because more than half of the sample participants were diagnosed with RD, it was unlikely that they start out at the same level of performance on reading and memory measures as do children without RD. Likewise, a profile of the data also showed a "hockey stick" pattern where children's performance decreased sharply or increased sharply. Furthermore, because we had only three data points at best, a curvilinear relation could not be calculated reliably. Thus, we used a piecewise linear regression model outlined in Laird and Ware (1982; for a discussion of this model, see also Littell, Milliken, Stroup, & Wolfinger, 1996, pp. 411–421). The model calls for two intercepts. One intercept determined intercept values for the total sample, and the other intercept focused on the classification variable (children without RD vs. children with RD). This piecewise linear random regression model is expressed as

$$y_{ij} = \beta_o + \beta_{o1}(\text{RD}) + \beta_1(\text{age}_{ij}) + U_{oj} + U_{02j} + U_{1j}(\text{age}_{ij}) + R_{ij},$$

where y_{ij} is the dependent variable (e.g., WM, reading comprehension, fluency) measured at time i in child j, age_{ij} is child j's age at time i, and β_o is the average intercept at 13.85. β_{o1} is the average deviation from that total sample value for children with and without RD. The variable was scored as a contrast variable (+3 for skilled readers and −1 for each subgroup of children with RD). A positive intercept in β_{o1} implied a larger increase in the level of performance for children without RD than from the sample as a whole. β_1 is the linear slope. U_{oj} is the random intercept for child j in the sample as a whole, and U_{02j} is the random intercept for the likeness of students in the same diagnostic category (children without and with RD). U_{1j} is the random age slope for child j. R_{ij} is the residual for child j at time i. The between-child variance components—$\tau_0^2 = \text{Var}(U_{0j})$, $\tau_{01}^2 = \text{Var}(U_{01j})$, and $\tau_1^2 = \text{Var}(U_{1j})$—reflected the individual differences in level of performance for the sample as a whole, the likeness of students within the diagnostic categories, and the rate of change between children, respectively . The fixed and random effect parameter estimates were obtained using PROC MIXED in SAS 9.1 (SAS Institute, 2003).

We estimated the association between the outcome (memory and reading) and repeated measures of age across the 3-year time periods. Both unconditional and conditional models are reported. For the unconditional model, reported are the fixed effects for the intercept value for Wave 3 at 13.85 years of age, the RD classification, and the average rate of change across individuals. For the random effects, the estimates are the variance around the sample intercept, the RD intercept, and the slope related to change over time. Significant random effects indicated that children differed in intercepts and/or rate of change (slopes). It is important to note that a nonsignificant random intercept emerged in most of our models for the RD classification variable, suggesting that not much variation existed "within" the two groups. That is, the random effect suggested that the "likeness" of the students was the same within the diagnostic category (RD or non-RD) (for a discussion of two-level random intercept models, see Snijders & Bosker, 1999, pp.64–65).

For the conditional model, we tested whether entering memory into the model as a reading predictor explained any statistically significant associations related to fixed and random effects. When one or more predictors are introduced into the conditional model, there are reductions in the magnitude of the various components. These reductions are analogous to effect sizes (Snijders & Bosker, 1999). This is similar to the use of R^2 in linear regression models. The primary distinction between a linear regression and HLM is that several R^2 values are relevant to HLM because there are several variance components.

Reliability estimates for the HLM model are based on the random effects. The random effect of the unconditional model represents the proportion of variance in that effect that is parameter specific rather than error variance. A random effect that has less than 5% parameter variance is considered unsatisfactory (Bryk & Raudenbush, 1992; Snijders & Bosker, 1999). If the random effect is significant and has satisfactory reliability, it is appropriate to test whether additional variables can explain some of the variance in the unconditional model. For example, as seen later in Table 5, the variance for the random slope is .27 and the residual is .10 when predicting reading comprehension in the unconditional model. The proportion of variance is 73% (.27/[.27 + .10]) and the slope is significant. Therefore, it is appropriate to test a conditional model that introduces the between-participant variables and tests the main effect relations of these variables. Fixed effect coefficients are used to test these effects. Specifically, the conditional model predicted that the reading level and slope would be associated with memory performance. To evaluate the compatibility of the data with our conditional model, we tested the significance of the model change. This was done by using the differences between the deviance value (i.e., lack of correspondence between model and data) from the unconditional and conditional growth models as chi-square values and the number of parameters that were added for the conditional model as degrees of freedom. A significant chi-square value indicated that the conditional model showed a better fit to the data than did the unconditional model.

Snijders and Bosker (1999) argued that the power to detect significant parameters in multilevel research is frequently low because of reductions in parameter reliability. For this reason, we maintained all multiple comparisons at $p < .05$.

Missingness

Missing data occur in longitudinal designs whenever an intended measurement is not obtained. Because we experienced some loss in our sample in the third year, we used a restricted maximum likelihood (REML) estimation to handle missing data (for a discussion, see Peugh & Enders, 2004). HLM allows for an incomplete data set, using data from all participants with at least two points, as it derives parameters and estimates in a two-step iterative manner. When compared with assumptions inherent in more traditional procedures (e.g., listwise and pairwise deletion), this procedure allowed us to keep cases with partial data while making less restrictive assumptions about the pattern of missingness (Schafer, 1997). Other procedures for handling missing data (listwise and pairwise deletion) required strong assumptions that data were missing completely at random. Thus, the presence or absence of an observed score must be completely unrelated to all other variables. This assumption was unlikely in our longitudinal study; therefore, the advantage of the REML estimation was that it provided valid estimates in spite of nonrandomly missing data for our small sample size (for a discussion of REML and sample size, see Snijders & Bosker, 1999, pp. 52–53, 175). Nevertheless, it is conceivable that missingness in our study also depended on unmeasured variables. To the extent that these unmeasured characteristics correlated with effects included in our analysis, the REML estimation was no guarantee against biased estimates. Regardless of this scenario, however, the estimates from REML estimation still were more accurate than estimates from more traditional methods such as listwise and pairwise deletion (e.g., Peugh & Enders, 2004).

Procedures

Four doctoral-level graduate students trained in test administration tested all participants in their school. One session of approximately 30 min was required for small group test administration, and three sessions of 45–60 min were necessary for individual administration. During the group testing session, data were obtained from the Raven's Colored Progressive Matrices test and the arithmetic subtest of the WRAT-3. The remaining tasks were administered individually. Test administration was counterbalanced to control for order effects.

Results

Appendixes A, B, and C show the descriptive data for each variable for all 3 years of the study as a function of risk group. The three waves are presented in terms of number of participating students, means, and standard deviations. Because there was attrition of the participants between Year 1 and Year 3, we compared the performance of the attrition group (children who did not participate in Year 3) and nonattrition (children who participated in all 3 years) on mean scores. The means and standard deviations between these two samples are shown in Table 1. ANOVAs were computed between the two groups (retained vs. not retained). No significant group differences emerged between the groups when a Bonferroni correction was made for Type I errors.

Multivariate analyses of variance (Wave 1)

Prior to our analysis of growth, a series of one-way multivariate analyses of variance (MANOVAs) were performed to examine differences in means on measures of STM, WM, comprehension, and fluency for Wave 1 data. The intercorrelations among the raw scores for the dependent measures used in the MANOVA are shown in Appendix D. Appendixes E and F show intercorrelations for Waves 2 and 3.

Table 1
Means and standard deviations of retained and nonretained participants

	Retained ($n = 51$)		Nonretained ($n = 32$)		*F* ratio	η^2
	M	*SD*	*M*	*SD*		
Age	11.86	2.53	12.19	2.64	0.33	.004
Verbal IQ	98.76	22.60	101.27	24.92	0.23	.002
Reading						
WTR2	86.75	13.18	92.87	18.52	3.06	.04
WTR1	32.64	6.79	35.42	8.31	2.80	.03
Arithmetic						
WTA2	94.83	15.32	96.65	17.11	0.25	.003
WTA1	32.16	5.45	33.71	6.03	1.47	.22
Comprehension						
PC2	84.64	14.18	85.84	22.50	0.09	.001
PC1	34.79	10.16	37.21	17.82	0.61	.007
Fluency						
SWE2	87.56	13.34	89.65	19.26	0.34	.004
SWE1	58.72	19.09	62.56	21.43	0.71	.008
Fluid intelligence						
RVN2	99.08	17.88	103.71	17.32	1.34	.02
RVN1	29.34	4.74	30.09	6.22	0.39	.004
Short-term memory						
PM	2.74	0.79	2.60	0.89	0.55	.006
WS	3.90	0.83	4.09	0.94	0.93	.01
DGSF	6.73	1.89	8.00	2.39	6.90*	.08
Working memory						
DGSB	4.18	1.76	4.61	2.78	0.71	.009
RYI	1.15	0.88	1.63	0.89	5.87*	.06
ADI	1.42	0.95	1.96	1.51	3.95	.05
Update	2.92	3.23	3.87	3.59	1.61	.02

Note. 1, standard score; 2, raw score; WTR, WRAT-3 reading; WTA, WRAT-3 arithmetic; PC, WRMT-R passage comprehension; SWE, TOWRE real words; RVN, Raven's; PM, phonological memory; WS, word span; DGSF, forward digit span; DGSB, backward digit span; RYI, rhyming; ADI, digit/sentence span; Update, updating.
* $p < .05$.

Short-term memory

A one-way MANOVA was conducted on span scores of phonological STM (pseudoword span, word span, and WISC-III digit span forward) and revealed a significant main effect for reading group, Wilks's $\lambda = .71$, $F(9, 170) = 2.84$, $p < 001$, $\eta^2 = .29$. The ANOVAs were significant for word span, $F(3, 75) = 4.09$,$p < .01$, $\eta^2 = .15$, and for digit span forward, $F(3, 75) = 4.01$, $p < .01$, $\eta^2 = .14$, but not for phonological memory, $F(3, 75) = 2.09$, $p > .01$, $\eta^2 = .08$. For the word span and digit span forward tasks, a Tukey's test showed that skilled readers outperformed children with RD only, the comorbid group, and low verbal IQ readers. The differences among subgroups of children with RD were not significant (children with RD only = comorbid group = low verbal IQ readers).

Working memory

A one-way MANOVA was also performed to examine reading group differences on measures of WM span scores. The result was significant, Wilks's $\lambda = .71$, $F(12, 182) = 2.08$, $p < .05$, $\eta^2 = .29$. Univariates were significant for backward digit span, $F(3, 75) = 3.60$, $p < .01$, $\eta^2 = .13$, for digit/sentence span, $F(3, 75) = 3.26$, $p < .05$, $\eta^2 = .12$, for updating, $F(3, 75) = 3.28$, $p < .05$, $\eta^2 = .12$, and for rhyming, $F(3, 75) = 3.58$, $p < .05$, $\eta^2 = .13$. A Tukey's test showed that skilled readers outperformed the children with RD only, the comorbid group, and low verbal IQ readers on the backward digit span, digit/sentence span, updating, and rhyming tasks. The scores among children with RD only, the comorbid group, and low verbal IQ readers were statistically comparable, ps > .05.

Reading comprehension and fluency

A one-way ANOVA was significant for the passage comprehension subtest from the WRMT-R, $F(3, 76) = 12.49$, $p < .001$, $\eta^2 = 35$. A Tukey's test showed significant differences (ps < .05) in favor of skilled readers (skilled readers > comorbid group = children with RD only > low verbal IQ readers).

Fluency

A one-way ANOVA was significant for real word reading efficiency, $F(3, 76) = 21.03$, $p < .0001$, $\eta^2 = .46$. A Tukey's test showed that skilled readers outperformed children with RD only, the comorbid group, and low verbal IQ readers on the fluency tasks. No significant differences were found among scores for children with RD only, the comorbid group, and low verbal IQ readers, ps > .05.

Data reduction

Prior to the analysis of growth, it was necessary to reduce the number of memory measures because of the sample size. All raw score measures were scaled to have a mean of 0 and a standard deviation of 1 at Wave 1. Waves 2 and 3 measures were z scores based on the means and standard deviations of Wave 1. It was necessary to scale to z scores across the total sample so that all parameters were on the same metric, enabling meaningful comparisons for both age and time (for a discussion, see McGaw & Jöreskog, 1971). Because the sample size was larger in Wave 1 than in the subsequent waves, we tested whether the component structure of the memory measures in Wave 1 reflected two constructs: STM and WM. To address this issue, we ran a confirmatory factor analysis using the CALIS (covariance analysis and linear structural equation) software program (SAS, 1992), with the four WM tasks (backward digit span, updating, digit/sentence span, and rhyming) loading on one factor and the three STM tasks (forward digit span, pseudoword span, and word span) loading on a separate factor. The fit statistics were .93 for the Comparative Fit Index (CFI) (Bentler & Wu, 1995) and .091 for the root mean square error of approximation (RMSEA) (Jöreskog & Sörbom, 1984), indicating a good fit to the data. All standardized parameters were significant at the .01 level. Because we had no theoretical reasons for relying on a one-factor model, the two-factor model was accepted as a good fit to the data for the total sample. Thus, the unique variance related to WM and STM composite scores was assessed in the next series of analyses. Composite scores for STM (mean z scores of digit forward span, pseudoword span, and word span) and WM (mean of z scores for backward digit span, updating, digit/sentence span, and rhyming) were computed for each testing wave. Scores for reading comprehension (passage comprehension subtest from the WRMT-R) and fluency (real word and reading efficiency) were z scores based on Wave 1 performance.

Correlations

Prior to our analysis of growth, we correlated Wave 1 composite scores to Wave 3 composite scores. As shown in Table 2, composite scores for STM, WM, comprehension, and fluency in Wave 1 were compared with the same measures in Wave 3. Age at Wave 1 was also included in the analysis. As shown, the stability of the measures (correlation between the same measures) was significant for all memory composite and reading scores. STM in Wave 1 was weakly correlated with reading comprehension and fluency in Wave 3. In contrast, WM performance in Wave 1 was significantly correlated with comprehension and

Table 2
Correlations between composite scores for Waves 1 and 3 for total sample

Wave 3	*M*	*SD*	Age	Wave 1			
				STM1	WM1	Comp1	Fluency1
STM3	−.09	.76	.17	.45***	.54***	.33**	.40**
WM3	.36	.66	.28*	.48***	.64***	.58***	.52***
Comp3	.26	.55	.38*	.31*	.57***	.75***	.53***
Fluency3	.40	.85	.24	.23	.52***	.55***	.74***
M			143.45	.01	.13	−.06	−.01
SD			30.85	.84	.81	.84	1.00

Note. Age, age in months; STM1, short-term memory, Wave 1; WM1, working memory, Wave 1; Comp1, reading comprehension, Wave 1; Fluency1, fluency, Wave 1; STM3, short-term memory, Wave 3; WM3, working memory, Wave 3; Comp3, reading comprehension, Wave 3; Fluency3, fluency, Wave 3.

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

fluency in Wave 3. Interestingly, correlations between the composite scores and age were in the low range (.17–.38). Thus, although age at Wave 1 was positively correlated with memory and reading performance, the magnitude of the correlations was marginal for the age groups represented in this study.

Growth among subgroups

Prior to modeling whether WM contributed to growth in reading, we used a mixed model with repeated measures to determine whether significant differences emerged in the level of performance among ability groups across testing waves. Table 3 shows the estimated parameters for each testing wave and whether estimates were greater than chance.

Short-term memory

A 4 (Ability Group: skilled readers, children with RD only, comorbid group, or low verbal IQ readers) × 3 (Testing Wave: 1, 2, or 3) repeated measures analysis, with chronological age (age at Wave 1) as the covariate, was computed on STM composite scores. A significant F ratio emerged for ability group, $F(3, 77) = 5.43$, $p < .01$, and for testing wave, $F(2, 107) = 5.11$. $p < .01$. No significant Ability Group × Testing Wave interaction occurred, $F(6, 107) = 1.97$, $p > .05$. The covariate was significant, $F(1, 77) = 10.50$, $p < .01$. The between-participant variance was significant (estimate = 22, $SE = .11$, z value = 1.94, $p < .05$), as was the within-participant variance (estimate = .36, $SE = .10$,

Table 3
Comparison of ability groups on intercept estimates of performance for STM, WM, reading comprehension, and fluency partialed for age at Wave 1

	STM		WM		Comprehension		Fluency	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Children with RD only								
Wave 1	−.31	.18	−.18	.15	−.33*	.14	−1.01**	.19
Wave 2	.31	.19	.23	.15	.22	.14	−.44*	.19
Wave 3	.07	.20	.39*	.16	.40*	.15	−.47*	.20
Comorbid group								
Wave 1	−.16	.19	−.26	.16	−.47*	.14	−1.26**	.19
Wave 2	.22	.22	−.16	.18	−.39*	16	−1.10**	.21
Wave 3	−.02	.25	.20	.20	.12	.12	−.91**	.23
Skilled readers								
Wave 1	.46**	.16	.75**	.13	.62**	.12	.23	.16
Wave 2	.70**	.17	.68**	.14	.82**	.12	.42**	.17
Wave 3	.38*	.20	.95**	.16	.86**	.14	.82**	.19
Low verbal IQ readers								
Wave 1	−.12	.16	−.04	.13	−.41**	.12	−.68**	.16
Wave 2	−.31	.17	.12	.14	−.09	13	−.51**	.17
Wave 3	.37	.19	.14	.15	−.16	.14	−.45**	.18

Note. Significant parameter estimates reflected scores that were better than chance.
* $p < .05$.
** $p < .01$.

z value = 3.37, $p < .001$). The estimated parameters for intercepts as a function of ability group and testing wave are shown in Table 3. A Tukey's test indicated that skilled readers outperformed (ps < .05) the other ability groups (skilled readers > children with RD only = comorbid group > low verbal IQ readers), and the estimated parameters were higher for Wave 2 than for Waves 1 and 3 (Wave 2 > Wave 1 = Wave 3).

Working memory

A 4 (Ability Group) × 3 (Testing Wave) repeated measures analysis, with chronological age (age at Wave 1) as the covariate, was computed for WM. A significant F ratio emerged for ability group, $F(3,77) = 9.95$, $p < .001$, and for testing wave, $F(2,107) = 9.10$, $p < .001$. No significant Ability Group × Testing Wave interaction occurred, $F(6,107) = 1.32$, $p > .05$. The covariate was significant, $F(1,77) = 20.13$, $p <$ 001. The between-participant variance was significant (estimate = 21, SE = .05, z value = 3.99, $p < .001$), as was the within-participant variance (estimate = .18, SE = 03, $z$value = 5.82, $p < .005$). A Tukey's test indicated that skilled readers outperformed (ps < .05) the other ability groups (skilled readers > children with RD only = comorbid group > low verbal IQ readers), and the parameters were higher for Waves 3 and 2 than for Wave 1 (Wave 3 = Wave 2 > Wave 1).

Reading comprehension

A 4 (Ability Group) × 3 (Testing Wave) repeated measures analysis, with chronological age (age at Wave 1) as the covariate, was computed for reading comprehension (passage comprehension). A significant F ratio emerged for ability group, $F(3,76) = 16.63$, $p < .001$, for testing wave, $F(2,103) = 26.74$, $p < .0001$, and for the Ability Group × Testing Wave interaction, $F(6,103) = 3.62$, $p < .001$. The covariate was significant, $F(1,76) = 50.92$, $p < .0001$. The between-participant variance was significant (estimate = 35.71, SE = 12.09, z value = 2.97, $p < .01$), as was the within-participant variance (estimate = 22.96, SE = 9.04, z value = 2.54, $p < .005$). A Tukey's test indicated that no significant differences emerged across testing waves (Wave 1 = Wave 2 = Wave 3) for low verbal IQ readers, whereas significant differences emerged across testing waves for skilled readers (Wave 1 < Wave 2 = Wave 3), for children with RD only (Wave 1 < Wave 2 < Wave 3), and for the comorbid group (Wave 1 = Wave 2 < Wave 3). Skilled readers outperformed the less skilled reading groups in Wave 1 (skilled readers > children with RD only = comorbid group > low verbal IQ readers), in Wave 2 (skilled readers > children with RD only > comorbid group = low verbal IQ readers), and in Wave 3 (skilled readers > children with RD only = comorbid group > low verbal IQ readers).

Fluency

A 4 (Ability Group) × 3 (Testing Wave) repeated measures analysis, with chronological age (age at Wave 1) as the covariate, was computed for fluency scores. A significant F ratio emerged for ability group, $F(3,76) = 18.52$, $p < .001$, and for testing wave, $F(2,104) = 12.82$, $p < .001$. No significant Ability Group × Testing Wave interaction occurred, $F(6,104) = 1.44$, $p > .05$. The covariate was significant, $F(1,76) = 17.35$, $p < .001$. The between-participant variance was significant (estimate = .35, SE = .09, z value = 3.81, $p < .001$), as was the within-participant variance (estimate = .21, SE = .05,

z value = 3.64, $p < 005$). A Tukey's test indicated that skilled readers outperformed ($ps < .05$) the other ability groups (skilled readers > children with RD only = low verbal IQ readers > comorbid group), and parameters were higher for Wave 3 than for Waves 1 and 2 (Wave 3 > Wave 2 > Wave 1).

In general, the results support the notion that skilled readers' levels of performance across all three testing waves on measures of STM, WM, reading comprehension, and fluency exceeded those of less skilled readers. Except for the reading comprehension measures, no significant Ability Group × Testing Wave interaction emerged, suggesting that the rates of growth were comparable among ability groups on memory measures and fluency.

Contrasts among subgroups

The purpose of the next analysis was to directly compare the intercept and growth differences between children with RD only and the comorbid group on measures of memory. Prior to this comparison, however, it was necessary to determine whether individual differences emerged in intercept and growth parameters. Table 4 shows the unconditional means growth model (Model 1) for the total sample. Presented in the table are the random portions (random effects) and overall centered means (fixed effects) for STM and WM. For STM, the estimated variance (random effects) of individual deviations from the overall intercept (.21) and slope (.03) was significantly different from zero. This result suggested that significant individual differences emerged in intercept and growth estimates of STM. Also shown in Table 4 is the residual (.33) that reflected the within-participant variability across testing waves. The table also shows the fixed effects. For STM, the fixed effect for the intercept indicated that the average z score across children at 13.85 years of age was .29 at Wave 3 and that the slope was .12. Hence, the average child in the total sample ended with a z score of .29 but gained .12 units per testing session. For WM, both the random and fixed effects were significant. As shown in Table 4, the average child at 13.85 years of age in the total sample yielded an estimate of .40 and gained .14 units per testing session. The classification variable was significant, indicating that skilled readers had greater STM and WM scores than did less skilled readers.

The conditional model shown at the bottom of Table 4 compared the children with RD only with the comorbid group on measures of memory. The model included the addition of an "RD-only vs. RD + AD" contrast variable (where AD is arithmetic deficit), age at Wave 1, fluid intelligence, and growth parameters related to the RD classification variable and the contrast variable. The contrast variable (RD-only was coded as −1, RD + MD was coded as +1, and low verbal IQ readers and skilled readers were coded as 0 [for a discussion of contrast variables in regression models, see Cohen & Cohen, 1983]) was used to enhance power in the group comparisons as well as to test for predicted differences among the subgroups. Because significant differences on measures of fluid intelligence emerged and it could be argued that memory differences are merely an artifact of general intelligence, scores on the Raven's Progressive Matrices test were also entered into the analysis as a covariate.

When compared with the unconditional means model, the conditional model accounted for approximately 23% ([.21 − .16]/21) of the explainable variance between individuals on level of STM performance and for approximately 99% ([.03 − .0002]/.03) of the explain-

Table 4
Comparison of ability groups on memory with age and fluid intelligence as covariates

	STM		WM	
(a) Unconditional model				
Random effects	Variance	*SE*	Variance	*SE*
Intercept	.21*	.10	.30***	.10
Growth	.03*	.004	.01*	.001
RD classification	.005	.02	.02	.02
Residual	.33***	.05	.17***	.02
Fixed effects	Estimate	*SE*	Estimate	*SE*
Intercept	.29**	.10	.40***	.10
Growth	.12**	.03	.14***	.02
RD classification	.13**	.04	.18***	.03
(b) Conditional model				
Random effects	Variance	*SE*	Variance	*SE*
Intercept	.16	.10	.28**	.11
Growth	.0002**	.00002	.0004***	.00007
RD classification	.0002	.01	.01	.01
Residual	.35**	.05	.17***	.02
Fixed effects	Estimate	*SE*	Estimate	*SE*
Intercept	−.96*	.49	−.84*	.44
Growth	.06	.05	.14**	.04
RD classification	.14**	.04	.16**	.04
Explanatory variables				
RD vs. RD + AD-intercept	.03	.09	−.07	.09
RD classification-growth	.01	.01	.004	.01
RD vs. RD + AD-growth	−.001	.02	−.02	.02
Age, Wave 1	.06	.06	−.02	.04
Fluid intelligence	.01**	.004	.01**	.003

Note. RD classification, RD subgroups vs. skilled readers.
* $p < .05$.
** $p < .001$.
*** $p < .001$.

able variance between individuals in STM growth. Likewise, the conditional model accounted for approximately 7% ([.30 − .28]/.30) of the explainable variance between individuals on level of WM performance and for approximately 96% ([.01 − .0004]/.01) of the explainable variance between individuals in WM growth. The important finding for the conditional model, however, was that no significant effects emerged for the RD-only vs. RD + MD contrast variable on either intercept or growth estimates for STM or WM. When skilled readers were compared with the RD groups, significant differences in intercept estimates in favor of skilled readers emerged on both STM and WM measures. An additional finding was that fluid intelligence was significantly related to both STM and WM performance. In general, no support was found to suggest that children with RD are superior to the comorbid group on measures of WM.

Influence of memory measures on growth in reading

In the final analysis, we investigated whether measures of STM and WM were related to growth in reading in the total sample. For Table 5, we first presented the unconditional means growth model (Model 1) for the total sample. Presented in the table are the random portions (random effects) and overall centered means (fixed effects) for reading comprehension and fluency. For reading comprehension, the estimated variance (random effects) of individual deviations from the overall intercept (.26) and slope (.27) was significantly different from zero. Also shown in Table 5 is the residual (.10) that reflected the within-participant variability across testing waves. The table also shows the fixed effects. For reading comprehension, the fixed effect for the intercept indicated that the average *z* score across children at 13.85 years of age was .14 at Wave 3 and that the slope was .15. Hence, the average child in the total sample ended with a *z* score of .14, but gained .15 units per testing session. For fluency, both the random and fixed effects were significant. As shown in Table 5, the average child at 13.85 years of age in the total sample yielded an estimate of .32 and gained .18 units per testing session. The classification variable was significant, indicating that skilled readers had higher comprehension and fluency scores than did less skilled readers.

The conditional model for reading comprehension and fluency is presented in Table 6. This conditional model was computed with memory intercept and growth parameter estimates entered into the model to test for the contribution of WM and STM to reading. Our assumption was that entering these two variables together would allow the unique variance related to WM and STM to reflect controlled attention and phonological storage, respectively (cf. Engle et al., 1999). We also entered verbal intelligence to control for differences in verbal skills. Because fluency has been considered an important predictor of comprehension performance, this covariate was also entered into the model estimates of reading comprehension. Furthermore, to control for age, children's chronological age at Wave 1 was entered into the mixed model. As shown in Table 6 for reading comprehen-

Table 5
Growth model for reading comprehension and fluency: Unconditional linear growth model

	Reading comprehension		Fluency	
Random effects	Variance	*SE*	Variance	*SE*
Intercept (variance)	.26***	.08	.52***	.14
Growth	.27*	.14	.24**	.08
RD classification	.16***	.06	.06	.04
Residual	.10***	.01	.20***	.03
Fixed effects	Estimate	*SE*	Estimate	*SE*
Intercept	.14***	.06	.32**	.08
Growth	.15***	.02	.18***	.05
RD classification	.31**	.06	.40***	.02
	Fit statistic		Fit statistic	
Deviance	273.1		380.3	

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

Table 6
Growth model for reading comprehension and fluency with covariates: Conditional growth model

	Reading comprehension		Fluency	
Random effects	Variance	*SE*	Variance	*SE*
Intercept	.12*	.06	.40**	.10
Growth	.20**	.21	.21**	.002
RD classification	.02	.02	.05	.04
Residual	.10***	.01	.19**	.02
Fixed effects	Estimate	*SE*	Estimate	*SE*
Intercept	−.60**	.21	−.06	.35
Growth	.13**	.03	.17**	.04
RD classification	.10*	.05	.27**	.04
Explanatory variables				
Starting age at Wave 1	−.02	.03	−.05	.004
Verbal intelligence	.009**	.002	.003	.002
Fluency	.19**	.04	—	—
STM	.06	.04	.03	.06
WM	.09	.05	.22**	.07
Linear growth				
STM-growth	−.008	.01	−.02	.02
WM-growth	−.04**	.02	−.05*	.02
	Fit statistic		Fit statistic	
Deviance	247.1		377.9	

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

sion, the estimated variance in the intercept between individuals was .12 and the estimated variance in growth between individuals was 20. When compared with the unconditional model (Table 5), the conditional model accounted for approximately 53% ([26 − .12]/.26) of the explainable variance between individuals on level of performance and for approximately 25% ([.27 − .20]/.27) of the explainable variance between individuals in reading comprehension growth. The residual (estimated variance within participants across the three testing waves) remained stable between the two models. Parameters that were significant explanatory variables in the fixed effects portion of the reading comprehension model were intercept values of verbal intelligence, fluency, WM, and the WM linear growth parameter.

Unconditional and conditional models were computed to examine reading fluency. When compared with the unconditional model, the conditional model accounted for approximately 23% ([.52 − .40]/52) of the explainable variance between individuals on level of performance and approximately 13% ([.24 − .21]/.24) of the explainable variance between individuals in reading fluency growth. Parameters that contributed significantly to the fixed effects portion of the conditional fluency model were WM intercept and growth parameters. A deviance test indicated that the conditional model showed a better fit to the data than did the unconditional model for reading comprehension, $\chi^2(df = 7) = 26.00$ (273.1 − 247.1), $p < .001$, but not for fluency, $\chi^2(df = 6) = 2.40$ (380.3 − 377.9), $p > .05$.

As the reader can surmise, a number of nonsignificant parameters emerged in the conditional model for both reading comprehension and fluency. Thus, we investigated more parsimonious models that included only the significant effects. The reduced conditional models are shown in Table 7. These models were compared with the unconditional models (Table 5). A deviance test indicated that the reduced conditional model showed a better fit to the data than did the unconditional model for reading comprehension $\chi^2(df = 4) = 38.20$, $p < .001$, and for fluency, $\chi^2(df = 2) = 16.70$, $p < .001$ For reading comprehension, the reduced model accounted for 39% ([.26 − .16]/.26) of the explainable

Table 7
Growth model for reading comprehension and fluency with covariates: Reduced conditional growth model

	Reading comprehension		Fluency	
Random effects	Variance	*SE*	Variance	*SE*
Intercept	.16**	.05	.41**	.11
Growth	.22**	.06	.23**	.05
RD classification	.01	.03	.04	.03
Residual	.10***	.01	.19**	.03
Fixed effects	Estimate	*SE*	Estimate	*SE*
Intercept	−.68***	.19	.26**	.08
Growth	.13***	.01	.13***	.02
RD classification	.12*	.04	.30**	.05
Starting age at Wave 1	—	—	—	—
Verbal intelligence	.01**	.002	—	—
Fluency	.121*	.04	—	—
STM	—	—		
WM	—	—	.24**	.07
Linear growth				
STM	—	—		
WM	−.04**	.01	−.08**	.02
	Fit statistic		Fit statistic	
Deviance	234.9		363.6	
Reduced conditional growth model with WM only				
Random effects	Variance	*SE*		
Intercept (variance)	.11*	.08		
Growth	.11*	.05		
RD classification	.02	.06		
Residual	.10***	.02		
Fixed effects	Estimate	*SE*		
Intercept	.18*	.06		
Growth	.15***	.02		
RD classification	.27**	.05		
WM-linear growth	−.08**	.01		
	Fit statistic			
Deviance	258.4			

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

variance in intercept parameters and 19% ([.27 − .22]/.27) of the explainable variance in growth estimates. For reading fluency, the reduced model accounted for 21% ([.52 − .41]/.52) of the explainable variance in intercept parameters but only 4% ([.24 − .23]/.24) of the variance in growth estimates.

We reexamined the reduced model for reading comprehension by eliminating all of the significant explanatory variables except for the WM growth estimates. (A WM model is already reflected in the reduced model for fluency.) The estimates are reported at the bottom of Table 7. When compared with the unconditional model, the reduced WM model accounted for approximately 58% ([.26 − .11]/.26) of the explainable variance in the intercepts and 19% ([.27 − .22]/.27) of the explainable variance in growth for reading comprehension. A deviance test indicated that the reduced memory conditional model showed a better fit to the data than did the unconditional model for reading comprehension, $\chi^2(df = 1) = 14.70$, $p < .001$.

In summary, the important findings were that the estimates of growth and level of performance in WM were significantly related to reading comprehension when the influence of age, STM, fluency, and verbal IQ were partialed out in the analysis. In addition, the overall level of performance in WM, and not in STM, was significantly related to fluency.

Discussion

The purpose of this study was to determine (a) whether children with RD only outperformed children with comorbid deficits on measures of WM and (b) whether an executive system (controlled attention) and/or a phonological storage system account for growth in reading performance. In terms of the study's first purpose, the results show that the level of performance and growth on measures of WM are statistically comparable between the children with RD only and the comorbid group even when fluid intelligence and age are partialed from the analysis (Table 4). As expected, the results also show that memory growth for skilled readers differed significantly from that for subgroups of children with RD on measures of STM and WM (Table 4). In addition, the results related to the repeated measures show no significant differences among subgroups of children with RD in memory performance. In terms of the study's second purpose, growth modeling for the total sample shows that WM (controlled attention), rather than STM (phonological loop), was significantly related to growth in reading. We now consider two questions that motivated this study.

Do different growth patterns in WM emerge between children with RD and skilled readers and between children with RD only and children with more generalized academic problems?

Some studies (e.g., van der Sluis et al., 2005) have suggested that children with RD only exhibit no problems in executive function, whereas children who show reading plus arithmetic difficulties have problems in executive processing. Thus, of particular interest in our study was to determine whether the performance of children identified with only RD reflects STM deficits in the phonological system, whereas the performance of children with comorbid deficits in both reading and arithmetic reflects deficits in activities related to the executive system. We did not find support in the HLM analysis for the notion that more generalized WM deficits (controlled attention) and storage deficits occur only in the comorbid group when compared with children with RD only. Furthermore, because more generalized deficits in achievement (problems in both reading and arithmetic) may be associated with verbal IQ, we also explored whether group differences in WM measures were merely an artifact of verbal ability. No significant differences were found between low verbal IQ readers and children with RD only on measures of STM and WM. Thus, we did not find support for the notion that WM deficits in children with RD are merely an artifact of verbal IQ or comorbidity. These findings coincide with Swanson's (2003) cross-sectional study that included elementary and secondary school children. Swanson found no significant differences in span scores on verbal and visual–spatial WM across a broad age range in children with RD, suggesting that memory growth may be arrested in children with RD during the late elementary and secondary school years.

Because both STM and WM intercept estimates of children with RD significantly differed from those of skilled readers (Table 4), we infer that children with RD may suffer WM deficits that operate in parallel with STM deficits. How could this be? Problems in WM have been found to persist in children with RD even after partialing out the influence of verbal articulation speed (Swanson & Ashbaker, 2000), reading comprehension (Swanson, 1999), STM (Swanson et al., 1996), and IQ scores (Swanson & Sachse-Lee, 2001). However, there is abundant evidence that children with RD suffer deficits in STM, a substrate of the phonological system. Children with RD, when compared with skilled readers, perform poorly on tasks that require accurate recall of letter and number strings or real words and pseudowords. Tasks such as these, which have a "read in and read out" quality (i.e., placing few demands on LTM to infer or transform the information), reflect STM. One common link among these tasks is the ability to store and/or access the sound structure of language (phonological processing). Likewise, these same children with RD also do poorly on tasks that place demands on attentional capacity, a characteristic of WM tasks. These combined deficit findings make sense because there are many mnemonic situations where a stimulus in memory is attended to and the other stimuli exist as a background; that is, they are not the center of current awareness. WM tasks require the active monitoring of events, and these events appear to be distinguishable from simple attention to stimuli held in STM.

Does an executive system (controlled attention) and/or a phonological storage system account for growth in reading performance?

Our results show that growth in WM is essential for the mental activities assumed to be basic to children's high-level processing such as comprehension. The results show that for the reduced model that included only WM, approximately 58% of explainable participant variance in level of reading comprehension performance and 21% of explainable participant variance in level of fluency performance were explained by WM performance. More important, in the reduced model that included only WM, approximately 19% of explainable participant variance in reading comprehension growth and 4% of explainable participant variance in growth in fluency performance were explained by WM performance.

However, because a part of WM distinguishes between STM and controlled attention, the question arises as to whether WM as an attentional mechanism or WM as storage (STM) is an important source of children's growth in reading. Both STM and WM have been found in some studies to make contributions to reading problems in children with RD (e.g., de Jong, 1998; Passolunghi & Siegel, 2001; Siegel, 1994; Stanovich & Siegel, 1994; Swanson, 1994; Swanson & Ashbaker, 2000). In the current longitudinal study, we attempted to break out the contributions of WM by focusing on both STM and WM measures. In our interpretation of the findings, we draw heavily from the work of Engle and Cowan. That is, there is some consensus in the literature that WM and STM are distinct, but highly related, processes. For example, Engle and colleagues (1999) found that STM and WM tasks loaded onto two different factors. Although strong correlations emerged between the two factors ($r = .70$), the authors found that a two-factor model fit the data better than did a one-factor model. Furthermore, they found that by statistically controlling the variance between WM and STM factors, the residual variance related to the WM factor was significantly correlated with measures of fluid intelligence (e.g., Raven Progressive Matrices test, Cattell Culture Fair Test). That is, they found a strong link between the latent measures of WM, but not of STM, to fluid intelligence. Engle and colleagues interpreted their findings as suggesting that the residual variance related to the WM factor corresponded to controlled attention of the central executive system and that this system, in turn, was strongly correlated with measures of fluid intelligence. In elaborating the distinction between STM and WM, Cowan (1995) emphasized the role of attentional processing in children's WM development. WM is described as a subset of items of information that are stored in STM and submitted to controlled attentional processing. Consistent with these assumptions, our results show that when the total sample is examined, controlled attention is related to children's growth in reading. The important findings from our study are that the growth and level of performance in WM were significantly related to reading comprehension and reading fluency when STM was entered into the analysis.

What are the theoretical implications of our findings? Two are considered here. First, these results extend those studies on individual differences suggesting that a WM system plays a critical role in reading growth. In terms of individual differences, children who have a large WM capacity for language can carry out the execution of various fundamental reading processes with fewer demands on a limited resource pool than can children who have a smaller WM capacity. As a result, children with a larger WM capacity would have more resources available for storage while comprehending a passage. In contrast, children with a smaller WM capacity might have fewer resources available for the maintenance of information during reading. Furthermore, this relation holds (at least for children) even when the influence of phonological loop (STM) is partialed from the analysis. Yet WM is not the exclusive contributor to variance in reading growth. The study also supports previous research about the importance of fluency in reading comprehension.

Second, the results address two competing models of WM deficits in children with RD. One model suggests that the phonological loop (STM) plays a major role in literacy growth. The phonological loop in this study was related to composite scores of STM. The model follows logically from the reading literature that links phonological skills to new word learning (e.g., Baddeley, Gathercole, & Papagno, 1998) and comprehension (Perfetti, 1985). The model assumes that inefficiencies in the phonological loop provide a more parsimonious explanation of ability group differences in literacy growth than do measures of WM. The model suggests that children who are weak in literacy skills have deficits in the processing of phonological information, creating a bottleneck in the flow of information to higher levels of processing.

A second model suggests that growth in measures of literacy relates to executive processing independent of the influence of the phonological system. This assumption follows logically from the literature on comprehension and reasoning suggesting that high-level performance on these measures requires the coordination of several basic processes (e.g., Engle et al., 1999; Just et al., 1990; Kyllonen & Christal, 1990). Measures of executive processing in this study were related to WM. Our results show that controlled attention, rather than phonological storage, underlies subsequent reading growth. These findings are in line with Gathercole and colleagues' (2005) study with younger children in which general working memory skills appeared to constrain the acquisition of academic skills. Our data also provide evidence that the executive system plays an important role in overall outcomes on reading literacy.

There are, of course, several limitations to our study. Three are apparent. First, we focused on children who came from average- to high-SES homes with highly educated parents. We assumed that previous studies that have tested children with dyslexia had failed to control for SES and that, therefore, differences in reading could be attributed to environmental and/or instructional factors. Regardless, our findings can be generalized to only a select group of children with RD. A second limitation is that we failed to assess children on all components of Baddeley's WM model. Clearly missing from our battery of tests are measures that assess the episodic buffer and the visual–spatial sketchpad. At the time when we implemented the study, a consensus of research studies using measures to tap the episodic buffer were not available. A meta-analysis of the earlier literature suggests that visual storage for children with RD is intact (O'Shaughnessy & Swanson, 1998). However, these findings may be a function of processing demands (for a review, see Swanson, 2005, pp. 413–414). Therefore, our findings related to verbal measures do not imply that growth in the visual–spatial sketchpad is normal in children with RD. Finally, the reliability of individual tasks varied over the three testing waves. As shown in the appendixes, reliability coefficients were generally higher for the reading tasks than for the memory tasks. Although we greatly enhanced the reliability of memory analyses by relying on composite scores, caution in the interpretation of individual measures of memory is necessary.

Conclusion

Our results support the position that the constraints in the executive system play an important role when accounting for reading growth in skilled readers and children with RD. Children who are less skilled readers, regardless of their verbal intelligence or concurrent arithmetic difficulties, showed lower levels of memory performance than did skilled readers (Table 4). These memory skills were related to overall attainments in reading comprehension and fluency. The overall inferior achievements of children with RD, when compared with the achievements of skilled readers, result from pervasive memory deficits in the executive system that plays a primary role in literacy growth.

Acknowledgments

This research is part of a longitudinal study partially funded by Peloy Funds and by the U.S. Department of Education, Cognition, and Student Learning (USDE R305H020055), Institute of Education Sciences, awarded to the first author. The authors are indebted to Diana Dowds, Crystal Howard, and Leilani Sáez for the data collection. The authors are thankful to the Frostig school staff and teachers, especially Marshall Raskind and Eleanor

Higgins, for their administrative support. The authors also thank the three anonymous reviewers for their incisive comments on an earlier version of this chapter.

Appendix A

Means and standard deviations by ability group for Wave 1

	Reliability	Children with RD only ($n = 18$)		Comorbid group ($n = 18$)		Skilled readers ($n = 23$)		Low verbal IQ readers ($n = 25$)	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age		10.29	2.58	12.48	2.74	11.31	2.19	13.50	1.80
Verbal IQ	.93	100.86	16.85	106.41	17.79	116.60	27.40	78.64	3.61
Reading									
WTR2	.91	79.94	7.32	81.23	14.32	106.43	10.33	85.12	12.55
WTR1	.89	27.27	5.71	31.66	8.03	39.39	5.29	34.68	5.82
Arithmetic									
WTA2	.81	105.05	7.57	80.20	8.10	105.52	17.16	88.76	11.97
WTA1	.71	32.33	6.38	28.75	5.80	35.08	5.19	33.52	4.30
Comprehension									
PC2	.91	81.00	14.20	80.11	19.32	100.47	17.85	77.36	7.98
PPC1	.89	27.66	12.03	32.66	12.90	45.04	16.44	34.28	6.20
Fluency									
SWE2	.87	83.41	10.39	78.83	11.54	104.04	17.14	84.20	8.88
SWE1	.94	43.94	16.82	52.33	21.29	74.63	16.46	64.28	12.78
Fluid intelligence									
RVN2	.93	107.66	16.65	96.25	16.53	107.52	16.67	92.88	16.67
RVN1	.77	30.05	5.26	28.50	6.72	31.00	4.22	28.80	5.35
Short-term memory									
PM	.69	2.33	1.02	2.94	0.63	2.82	0.49	2.64	0.99
WS	.83	3.50	0.85	3.77	0.87	4.39	1.03	4.08	0.49
DGSF	.59	6.50	2.03	6.88	2.47	8.09	1.99	7.16	2.07
Working memory									
DGSB	.57	3.18	1.79	4.29	2.99	5.45	1.92	4.16	1.67
RYI	.76	0.72	0.82	1.33	1.08	1.82	0.77	1.36	0.70
ADI	.70	1.23	1.09	1.23	0.97	2.08	1.12	1.80	1.44
Update	.87	2.72	3.39	1.66	2.35	5.08	3.80	3.24	3.01

Note. 1, standard score; 2, raw score; WTR, WRAT-3 reading; WTA, WRAT-3 arithmetic; PC, WRMT-R passage comprehension; SWE, TOWRE real words; RVN, Raven's; PM, phonological memory; WS, word span; DGSF, forward digit span; DGSB, backward digit span; RYI, rhyming; ADI, digit/sentence span; Update, updating.

Appendix B

Means and standard deviations by ability group for Wave 2

	Reliability	Children with RD only ($n = 17$)		Comorbid group ($n = 14$)		Skilled readers ($n = 19$)		Low verbal IQ readers ($n = 20$)	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age		11.40	2.57	13.36	2.99	12.31	2.18	14.51	1.96
Verbal IQ	.87	97.28	13.33	86.66	13.06	100.06	14.04	82.00	7.62
Reading									
WTR2	.91	88.00	8.31	78.76	22.38	108.63	12.31	87.40	11.02
WTR1	.75	31.94	5.28	33.30	6.40	42.05	6.24	36.65	5.72
Arithmetic									
WTM2	.89	96.18	14.92	79.35	14.63	102.55	19.12	87.61	13.55
WTM1	.76	32.00	6.67	29.35	7.13	36.22	6.66	34.16	5.06
Comprehension									
PC2	.92	87.50	13.39	78.42	21.94	102.94	14.10	82.20	9.65
PC1	.89	34.93	9.81	33.35	14.39	44.26	10.02	38.55	5.08
Fluency									
SWE2	.88	88.37	9.97	78.46	13.21	101.00	17.64	83.47	11.07
SWE1	.94	59.12	11.95	54.07	23.23	75.78	15.58	66.65	13.35
Fluid intelligence									
RVN2	.88	112.68	8.89	98.23	14.34	106.05	15.09	94.77	13.32
RVN1	.68	32.68	2.52	29.46	4.94	32.44	3.60	29.94	5.34
Short-term memory									
PM	.47	2.82	1.13	2.85	0.77	2.94	0.70	2.15	1.13
WS	.70	3.82	0.95	3.92	0.99	4.15	1.21	3.95	0.68
DGSF	.95	7.76	2.41	8.07	2.12	9.47	2.52	7.70	2.07
Working memory									
DGSB	.57	4.29	1.79	4.00	2.11	4.84	1.64	4.55	1.93
RYI	.70	1.50	1.03	1.35	0.92	1.88	0.90	1.55	0.94
ADI	.80	1.25	0.93	1.35	0.84	2.17	1.13	1.70	1.03
Update	.95	2.94	2.65	3.00	3.00	4.10	3.44	3.50	2.78

Note. 1, standard score; 2, raw score; WTR, WRAT-3 reading; WTA, WRAT-3 arithmetic; PC, WRMT-R passage comprehension; SWE, TOWRE real words; RVN, Raven's; PM, phonological memory; WS, word span; DGSF, forward digit span; DGSB, backward digit span; RYI, rhyming; ADI, digit/sentence span; Update, updating.

Appendix C

Means and standard deviations by ability group for Wave 3

Variable	Reliability	Children with RD only ($n = 14$)		Comorbid group ($n = 10$)		Skilled readers ($n = 12$)		Low verbal IQ readers ($n = 15$)	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age		11.99	1.85	14.26	2.96	13.63	1.73	15.84	1.64
Verbal IQ	.78	88.92	11.06	83.42	4.54	94.00	10.26	80.00	4.57
Reading									
WTR2	.88	86.50	10.52	86.10	11.64	102.83	6.84	83.46	12.28
WTR1	.71	33.07	4.77	35.40	4.45	42.25	3.81	36.26	5.96
Arithmetic									
WTM2	.90	93.50	14.43	80.40	15.31	93.83	14.35	82.53	13.75
WTM1	.72	33.50	6.17	30.80	6.26	36.41	6.41	33.80	4.88
Comprehension									
PC2	.83	89.64	9.33	92.20	7.92	96.16	9.79	80.78	8.07
PC1	.85	36.64	8.27	42.60	8.08	43.91	8.03	38.14	5.95
Fluency									
SWE2	.89	87.07	10.92	77.90	9.25	103.66	15.12	80.13	10.79
SWE1	.93	60.85	13.35	59.80	15.98	84.91	13.61	68.60	14.09
Fluid intelligence									
RVN2	.92	112.71	17.63	98.30	16.87	102.66	10.83	96.93	20.35
RVN1	.79	33.28	2.16	30.40	3.80	32.50	2.19	30.53	5.38
Short-term memory									
PM	.62	2.35	0.84	2.80	0.63	2.58	0.90	2.15	1.06
WS	.34	3.71	0.72	3.70	0.67	4.16	0.93	3.69	0.85
DGSF	.58	7.07	1.32	7.20	1.98	8.25	2.05	7.80	2.11
Working memory									
DGSB	.65	4.14	1.35	4.40	1.71	4.50	1.00	4.33	1.49
RYI	.74	1.42	0.64	1.20	1.03	1.75	0.86	1.40	1.05
ADI	.51	1.78	1.12	2.40	0.96	2.91	1.31	1.86	0.91
Update	.93	3.64	2.61	3.70	3.49	5.16	3.53	3.69	3.90

Note. 1, standard score; 2, raw score; WTR, WRAT-3 reading; WTA, WRAT-3 arithmetic; PC, WRMT-R passage comprehension; SWE, TOWRE real words; RVN, Raven's; PM, phonological memory; WS, word span; DGSF, forward digit span; DGSB, backward digit span; RYI, rhyming; ADI, digit/sentence span; Update, updating.

Appendix D

Intercorrelations among raw score measures at Wave 1

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	—												
2. VIQ	−.22*	—											
3. WTR	.55***	.19	—										
4. PC	.52***	.30**	.82***	—									
5. SWE	.48***	.18	.86***	.80***	—								
6. RVN	.19	.21*	.37***	.47***	.32**	—							
7. PM	.20	.15	.22*	.30**	.20	.16	—						
8. WS	.17	.32**	.46***	.46***	.47***	21	.35**	—					
9. DGSF	.36***	.26*	.56***	.48***	.51***	38***	.24*	.41***	—				
10. DGSB	.37***	.19	.57***	.58***	.59***	36***	.38***	.43***	.56***	—			
11. RYI	.15	.25*	.39***	.41***	.38***	17	.21*	.34**	.19	.26*	—		
12. ADI	.29**	.15	.49***	.52***	.47***	29**	.06	.26*	.43***	.34**	.30**	—	
13. Update	.27*	.27*	.52***	.53***	.54***	35***	.20	.34**	.34**	.42***	.25*	.35**	—
M	143.45	99.31	33.78	34.40	60.52	29.34	2.71	4.01	7.20	4.27	1.38	1.67	3.40
SD	30.95	24.38	7.31	11.59	20.55	5.60	0.80	0.85	2.09	1.94	0.85	1.25	3.35

Note. VIQ, verbal IQ; WTR, WRAT-3 reading; PC, WRMT-R passage comprehension; SWE, TOWRE real words; RVN, Raven's; PM, pseudoword memory; WS, word span; DGSF, forward digit span; DGSB, backward digit span; RYI, rhyming; ADI, digit/sentence span; Update, updating.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Appendix E

Intercorrelations among raw score measures at Wave 2 ($N = 68$)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	—												
2. VIQ	−.40*	—											
3. WTR	.49***	.23	—										
4. PC	.42***	.27	.83***	—									
5. SWE	.30*	.32*	.70**	.65***	—								
6. RVN	.09	.21	.20	.40***	.26	—							
7. PM	−.11	.32**	.03	.10	.06	.14	—						
8. WS	.15	.12	.25	.34**	.20	.13	.09	—					
9. DGSF	.16	.40**	.50**	.35*	.43**	.17	.22	.30*	—				
10. DGSB	.17	.29*	.30*	.30*	.40**	.26	.08	.28*	.31**	—			
11. RYI	.04	.17	.21	.24*	.15	.26	.08	.28*	.31*	.22	—		
12. ADI	.10	.32*	.32*	.29*	.40**	.01	.28*	.27*	.37**	.20	.10	—	
13. Update	.36**	.07	.52***	.43**	.21	.05	.11	.38**	.30*	.31**	.29*	.21	—
M	154.6	91.35	36.29	37.95	64.71	31.22	2.67	3.95	8.25	4.41	1.59	1.64	3.35
SD	31.82	13.91	7.08	10.55	17.76	4.38	1.01	0.96	2.24	1.82	0.96	1.05	2.96

Note. VIQ, verbal IQ; WTR, WRAT-3 reading; PC, WRMT-R passage comprehension; SWE, TOWRE real words; RVN, Raven's; PM, pseudoword memory; WS, word span; DGSF, forward digit span; DGSB, backward digit span; RYI, rhyming; ADI, digit/sentence span; Update, updating.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Appendix F

Intercorrelations among raw score measures at Wave 3 ($N = 49$)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	—												
2. VIQ	−.14	—											
3. WTR	.35***	.30*	—										
4. PC	.33*	.46**	.54**	—									
5. SWE	.24	.33	.67***	.43***	—								
6. RVN	.07	.16	−.04	.15	.18	—							
7. PM	.17	.20	.17	.35*	.06	.16	—						
8. WS	.10	.16	.24	.02	.10	.13	.35*	—					
9. DGSF	.13	.14	.50**	.20	.27	.10	.13	.40**	—				
10. DGSB	.58*	.16	.41**	.33*	.18	.19	.24	.30*	.34*	—			
11. RYI	.16	.06	.22	.06	.21	.16	.30*	.40*	.14	.19	—		
12. ADI	.09	.25	.35*	.29	.27	.21	.04	.18	.13	.21	.18	—	
13. Update	.26	.07	.45***	.25	.32*	−.09	.16	.19	.12	.30*	.31*	.19	—
M	166.73	87.52	36.47	39.35	68.32	31.73	2.45	3.80	7.73	4.24	1.48	2.20	3.35
SD	29.73	10.16	5.83	7.53	17.14	3.79	0.90	0.82	1.86	1.21	0.89	1.15	2.92

Note. VIQ, verbal IQ; WTR, WRAT-3 reading; PC, WRMT-R passage comprehension; SWE, TOWRE real words; RVN, Raven's; PM, pseudoword memory; WS, word span; DGSF, forward digit span; DGSB, backward digit span; RYI, rhyming; ADI, digit/sentence span; Update, updating.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

References

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87*, 85–106.

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

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

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

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., & Logie, R. H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge, UK: Cambridge University Press.

Bentler, P. M., & Wu, E. J. C. (1995). *Structural equations program manual*. Encino, CA: Multivariate Software.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. London: Sage.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cowan, N. (1995). *Attention and memory: An integrated framework (Oxford Psychology Series, No. 26)*. New York: Oxford University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433.

de Jong, P. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology, 70*, 75–95.

Dempster, F. (1985). Short-term memory development in childhood and adolescence. In C. Brainerd & M. Pressley (Eds.), *Basic processes in memory* (pp. 209–248). New York: Springer.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972–992.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*, 309–331.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39*, 3–27.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*, 265–281.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*, 177–190.

Gathercole, S. E., Tiffany, C., Briscoe, J., & Thorn, A. (2005). Developmental consequences of poor phonological short-term memory function in childhood. *Journal of Child Psychology and Psychiatry, 46*, 598–611.

Henry, L. A., & Millar, S. (1993). Why does memory span improve with age? A review of the evidence for two current hypotheses. *European Journal of Cognitive Psychology, 5*, 241–287.

Jöreskog, K. G., & Sörbom, D. (1984). *LISREL VI: Analysis of linear structural relations by the method of maximum likelihood*. Mooresville, IN: Scientific Software.

Just, M. A., Carpenter, P. A., & Keller, T. A. (1996). The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review, 103*, 773–780.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence, 14*, 389–433.

Laird, N. M., & Ware, J. (1982). Random-effects models for longitudinal data. *Biometrics, 39*, 963–974.

Leather, C., & Henry, L. (1994). Working memory span and phonological awareness task as predictors of early reading ability. *Journal of Experimental Child Psychology, 58*, 88–111.

Littell, R. C., Henry, P. R., & Ammerman, C. B. (1998). Statistical analysis of repeated measures data using SAS procedures. *Journal of Animal Science, 76*, 1216–1231.

Littell, R. C., Milliken, W. W., Stroup, R. D., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.

McGaw, B., & Jöreskog, K. G. (1971). Factorial invariance of ability measures in groups differing in intelligence and social economic status. *British Journal of Mathematical and Statistical Psychology, 24*, 154–168.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of central executive. *British Journal of Psychology, 91*, 111–121.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*, 139–158.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Oberauer, K., Sub, H., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence, 31*, 167–193.

O'Shaughnessy, T., & Swanson, H. L. (1998). Do immediate memory deficits in students with learning disabilities in reading reflect a developmental lag or deficit? A selective meta-analysis of the literature. *Learning Disability Quarterly, 21*, 123–148.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80*, 44–57.

Perfetti, C. (1985). *Reading ability*. New York: Oxford University Press.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research, 74*, 525–556.

Raven, J. C. (1976). *Colored progressive matrices*. London: H.K. Lewis.

SAS. (1992). *Covariance analysis and linear structural equation* (CALIS) (SAS Technical Report No. P-229, Release 6.07). Cary, NC: Author.

SAS Institute. (2003). *SAS/STAT user's guide*. Cary, NC: Author.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Shankweiler, D., Crain, S., Katz, L., Fowler, A. E., Liberman, A. M., Brady, S. A., et al. (1995). Cognitive profiles of reading disabled children: Comparison of language skills in phonology, morphology, and syntax. *Psychological Science, 6*, 149–156.

Siegel, L. S. (1994). Working memory and reading: A life-span perspective. *International Journal of Behavioral Development, 17*, 109–124.

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

Singer, J. D. (2002). Fitting individual growth models using SAS PROC MIXED. In D. S. Moskowitz & S. Hershberger (Eds.), *Modeling intraindividual variability with repeated measures data: Methods and applications* (pp. 135–170). Mahwah, NJ: Lawrence Erlbaum.

Snijders, T. A., & Bosker, R. J. (1999). *Multilevel analysis*. CA: Sage: Thousand Oaks.

Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. *Developmental Review, 10*, 72–100.

Stanovich, K., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core difference model. *Journal of Educational Psychology, 86*, 24–53.

Stothard, S., & Hulme, C. (1992). Reading comprehension difficulties in children. *Reading and Writing: An Interdisciplinary Journal, 4*, 245–256.

Swanson, H. L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology, 84*, 473–488.

Swanson, H. L. (1994). Short-term memory and working memory: Do both contribute to our understanding of academic achievement in children and adults with learning disabilities? *Journal of Learning Disabilities, 27*, 34–50.

Swanson, H. L. (1995). *S-Cognitive Processing Test*. Austin, TX: Pro-Ed.

Swanson, H. L. (1999). Reading comprehension and working memory in skilled readers: Is the phonological loop more important than the executive system? *Journal of Experimental Child Psychology, 72*, 1–31.

Swanson, H. L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of Experimental Child Psychology, 85*, 1–31.

Swanson, H. L. (2005). Working memory, intelligence, and learning disabilities. In O. Wilhem & R. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 409–430). CA, Sage: Thousand Oaks.

Swanson, H. L., & Ashbaker, M. (2000). Working memory, short-term memory, articulation speed, word recognition, and reading comprehension in learning disabled readers: Executive and/or articulatory system? *Intelligence, 28*, 1–30.

Swanson, H. L., Ashbaker, M., & Lee, C. (1996). Learning disabled readers' working memory as a function of processing demands. *Journal of Experimental Child Psychology, 61*, 242–275.

Swanson, H. L., & Berninger, V. (1995). The role of working memory and STM in skilled and less skilled readers' word recognition and comprehension. *Intelligence, 21*, 83–108.

Swanson, H. L., & Sachse-Lee, C. (2001). A subgroup analysis of working memory in children with reading disabilities: Domain-general or domain-specific deficiency? *Journal of Learning Disabilities, 34*, 249–263.

Swanson, H. L., & Siegel, L. (2001). Learning disabilities as a working memory deficit. *Issues in Education, 7*, 1–48.

van der Sluis, S., van der Leij, A., & de Jong, P. F. (2005). Working memory in Dutch children with reading- and arithmetic-related LD. *Journal of Learning Disabilities, 38*, 207–221.

Wagner, R., & Torgesen, J. (1999). *Test of Word Reading Efficiency*. Austin, TX: Pro-Ed.

Wilkinson, G. S. (1993). *The Wide Range Achievement Test-3*. Wilmington, DE: Wide Range.

Willcutt, E. G., Pennington, B. F., Olson, R. K., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: In search of the common deficit. *Developmental Neuropsychology, 27*, 35–78.

Woodcock, R. W. (1998). *Woodcock reading mastery test-revised.* Minneapolis, MN: American Guidance.