Published in final edited form as:
Read Res Q. 2006 October 1; 41(4): 496–522. doi:10.1598/RRQ.41.4.4.

Becoming a fluent and automatic reader in the early elementary school years

PAULA J. SCHWANENFLUGEL,
University of Georgia, Athens, USA

ELIZABETH B. MEISINGER,
University of Georgia, Athens, USA

JOSEPH M. WISENBAKER,
University of Georgia, Athens, USA

MELANIE R. KUHN,
Rutgers University, New Brunswick, New Jersey, USA

GREGORY P. STRAUSS, and
University of Nevada–Las Vegas, USA

ROBIN D. MORRIS
Georgia State University, Atlanta, USA

Abstract

The goals of this study were to (a) develop an empirically based model regarding the development of fluent and automatic reading in the early elementary school years and (b) determine whether fluent text-reading skills provided benefits for reading comprehension beyond those accounted for by fluent word decoding. First-, second-, and third-grade children completed a series of reading tasks targeting word and nonword processing, text reading, spelling knowledge, autonomous reading, and reading comprehension. Structural equation modeling was carried out to evaluate how these skills operated together to produce fluent text reading and good comprehension. Evidence supported a *simple reading fluency model* for the early elementary school years suggesting that fluent word and text reading operate together with autonomous reading to produce good comprehension.

The development of fluent and automatic reading skills is considered a primary educational goal for elementary school–age children. Although there is no single definition of reading fluency, there is general agreement that fluent reading incorporates the ability to read quickly, accurately, and, when oral reading is considered, with expression (National Institute of Child Health and Human Development [NICHD], 2000). In fact, a recent study indicates that second- and third-grade children who read quickly and accurately also tend to read with expression, suggesting that this definition of reading fluency is quite appropriate (Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004). In general, fluent reading emerges in most children between first and third grade, when decoding skills are confirmed through practice (Kuhn & Stahl, 2003). Reading fluency, as measured either by quick and accurate word reading or by text-reading rate, continues to develop beyond this period (Hasbrouck & Tindal, 1992; Horn & Manis, 1987).

The development of reading fluency is viewed as important because of its relationship with improved comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001). LaBerge and Samuels's (1974) now-classic article presenting an automaticity theory of reading argued that proficient word-recognition skills underlie fluent reading and adequate comprehension of text. According to the model, fluent readers are characterized by the ability to read quickly and without conscious effort (Logan, 1997). Dysfluent beginning readers, by contrast, are identified by their excessively slow, laborious reading, which, in turn, impairs comprehension.

Automaticity is an important component of skilled reading. Whether considered in terms of reading or any other skill, automaticity is identified by a number of characteristics, some of which concern us here. The first of these are speed and accuracy, which seem to emerge simultaneously with practice (Logan, 1988). Practice strengthens connections between word and letter patterns in long-term memory (e.g., LaBerge & Samuels, 1974), unitizes these letter patterns in memory so that they can be processed as whole units (Anderson, 1987), and proliferates the availability of instances of these word and letter unit representations in long-term memory with every encounter of them during reading (Logan, 1997).

Another characteristic of automatic skill is autonomy, or the ability to initiate a task without actively attending to it. The automatic reader cannot help but process print, even when he or she may intend to avoid doing so. Skilled readers find themselves reading the conflicting “crawl” at the bottom of a television news story, even when they try to pay attention to the main story (Bergen, Grimes, & Potter, 2005). However, the automatic nature of skilled word recognition is usually helpful because it orients our attention to the reading. Occasionally, as noted above, this can get in the way of our goals. Experimentally, the autonomous nature of automatic word reading has been measured through the use of the Stroop task (Stroop, 1935) and its variants. In the Stroop task, the reader is typically asked to name a picture or color without reading distracting words that are embedded. Autonomy is indicated by the interference that the distracting print causes to the picture or color naming. Thus, a child who is an automatic reader will be slowed in the naming a picture of an apple with the word *doll* written on it compared with naming a picture without the distracting print.

Finally, automaticity allows cognitive resources to be used to benefit larger goals. For reading, as automaticity develops for skills like word recognition, attentional resources become available for comprehension. Thus, children move from relying on the slow letter-by-letter (or unit-by-unit) decoding to painlessly retrieving cued words from long-term memory (Logan, 1997). This means that children who have efficient word-recognition skills should be able to read connected text fluently and better understand what they read.

The three aspects of automaticity (speed, autonomy, and resource use) are thought to develop concomitantly with one another, but how closely their development coincides is uncertain (Logan, 1985; Paap & Ogden, 1981). Clearly, at the very beginning of skill development, word reading is extremely slow, very effortful, and resource intensive (i.e., requiring a great deal of attention). Once a child has attained a high level of skill, word and text reading is very quick, occurs without intention, and ensures that plenty of resources are left available for comprehension. In the middle of skill development, it is unclear how the autonomy feature of automaticity relates to speed and resource aspects. Thus, it is uncertain whether the extent to which young readers experience Stroop interference, for example, is relative to their reading fluency or to improved comprehension.

How quickly automaticity develops depends on the underlying mechanisms and how they may work. For example, in some views, this buildup to automatic processing is seen as slow and laborious, occurring over hundreds of decoding attempts (Schneider, Dumais, & Shiffrin, 1984), whereas in others automaticity can develop after correctly recognizing a word only once or twice (Logan, 1997; Logan & Klapp, 1991). If the latter is true, then Stroop interference should develop quickly as children develop the ability to read. Moreover, the benefits for reading fluency and comprehension might emerge even for very young readers whose word recognition during reading is better than that of their peers. Despite the popularity of this automaticity account of the development of reading skill, model building to assess how these aspects of automaticity in early reading operate together is rare.

The purposes of the current study were two. First, we wished to provide a more comprehensive account regarding how fluency, autonomy, and freed resources (in terms of improved comprehension) operate together in early reading skill development. Thus, children completed reading tasks targeted at each aspect of automaticity, and the interrelationships among them were examined.

Second, we wished to determine the relative adequacy of two potential variants of this automaticity view of reading fluency, depicted in Figures 1a and 1b. In one view, which we call the *text reading as mediator model*, text-reading fluency is seen as uniquely relevant for comprehension (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003; Kuhn & Stahl, 2003) in contrast with other indicators of reading fluency. Young readers may or may not execute a variety of cognitive processes relevant to the comprehension of text if they have the cognitive resources available to generate them, such as inferencing to fill in the gaps in the text, creating a mental model of the world described by the text, and processing sentence grammar. At minimum, additional cognitive resource benefits may derive from fluent text reading through the contextual activation of word meanings (Stanovich, 1980) and the automatic linguistic segmentation of text into major syntactic units (Young & Bowers, 1995). This view is supported by studies showing that the fluent reading of text accounts for additional variance in reading comprehension beyond that accounted for by isolated word reading alone (Jenkins et al., 2003). In Figure 1a, this partial mediating effect of text reading is shown by the indirect path leading from GORT–3 Text-Reading Rate and to WIAT Reading Comprehension.

Another view, which we call the *simple reading fluency model*, states that reading fluency is a general skill that includes skills related to word-reading fluency as well as those related to text-reading fluency. In this model, it is assumed that, early in the development of reading skills, word-recognition skills are the limiting factor in reading comprehension and that the cognitive resources gained from automaticity derive mainly from quick and accurate word reading. Moreover, the skills that allow children to coordinate the benefits from fluent word and text reading may not yet be fully developed. This view is supported by studies showing that reading skill in young elementary school readers can be described as a single factor consisting of word-recognition skill, text-reading fluency, and reading comprehension (Shinn, Good, Knutson, Tilly, & Collins, 1992). Further, the texts that children read increase in complexity as they proceed in school. These changes in complexity may include, among others, greater syntactic complexity (Harber, 1979) and fewer repetitions of specific words within texts (Hiebert & Fisher, 2005). If text complexity increases faster than children's text fluency resources can be expended to capitalize on it, children may not experience the benefits gained from fluent text reading to support increases in comprehension during the early stages of fluent reading. Thus, children's word-recognition skills may need to become yet more automatic to manage these more complex text features given the resources children have available. In Figure 1b, this general influence of word-reading fluency is illustrated by showing GORT–3 Text-Reading Rate as merely another indicator of the latent variable Reading Fluency.

In what follows, we describe our rationale for including specific tasks and assessments as indicators of model features in tests of these larger models. We begin with our discussion of tasks and assessments that should be most closely related to reading fluency per se. Then, we describe research on tasks designed to assess the autonomy component and our assessment of reading comprehension.

Reading fluency component

Word-reading fluency

It is often asserted that the development of efficient word-recognition skills is the *sine qua non* of skilled reading in the early stages of learning to read (Adams, 1990; Lyon, Shaywitz, & Shaywitz, 2003; NICHD, 2000; Olson, Gillis, Rack, DeFries, & Fulker, 1991; Stanovich, 1985; Stanovich, Nathan, & Vala-Rossi, 1986; Stanovich, Nathan, & Zolman, 1988). It is the core skill around which reading fluency is built and an important skill for predicting reading comprehension (Gough, 1996; Perfetti & Hogaboam, 1975; Schwanenflugel et al., 2004). It distinguishes skilled from less skilled readers in a manner that is remarkably stable from first to fourth grade (Juel, 1988). It is often considered one, if not the key, bottleneck on the way to good reading comprehension (LaBerge & Samuels, 1974; Lyon, 1995; Nicholson, 1999; Perfetti, 1985).

A number of word features seem to enter into the skilled reading of words in young children. High-frequency words tend to be named faster than low-frequency words (Gottardo, Chiappe, Siegel, & Stanovich, 1999; Waters, Seidenberg, & Bruck, 1984). Orthographically irregular words (words whose spellings do not follow typical letter–sound correspondences) tend to be read more slowly and inaccurately than orthographically regular words in young readers (Backman, Bruck, Hebert, & Seidenberg, 1984; Nation & Snowling, 1998; Waters et al., 1984). Semantically difficult words (words rated low in imageability; Laing & Hulme, 1999; Nation & Snowling; Schwanenflugel & Akin, 1994; or rated context availability; McFalls, Schwanenflugel, & Stahl, 1996; Schwanenflugel & Noyes, 1996) are read more slowly and inaccurately than semantically easy words in young readers.

In the current study, we assessed children’s word-reading fluency in two ways. First, we assessed children’s automaticity for sight words using the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (TORP; 1999), a speeded sight-word–reading test. Second, we assessed children’s speed of processing for a potentially wider range of word types by presenting them with a mixed variety of words in a timed computerized single-word–naming task.

Nonword reading

Nonword reading is often used as an indicator of children’s understanding of the relationships between letters, letter strings, and larger units such as rime units and speech sounds (Coltheart & Leahy, 1992; Treiman, Goswami, & Bruck, 1990). Nonword reading is used as a way of assessing children’s ability to read unfamiliar words (Gottardo et al., 1999; Rack, Snowling, & Olson, 1992). Nonword-reading speed is widely considered an indicator of the automaticity of essential phonics and blending skills and is usually highly correlated with word-recognition skills (Torgesen, Wagner, & Rashotte, 1999). In the current study, we assessed children’s nonword-reading skills using the Phonemic Decoding Efficiency subtest from the Test of Word Reading Efficiency (Torgesen et al.), a speeded nonword-naming task.

Rapid object naming

Assessment of rapid object naming involves measuring how rapidly a child can carry out rapid naming of series of pictures. Although not obviously related to reading fluency, the speed with which children can carry out the rapid naming of pictures, as well as numbers, letters, and colors, has been shown to be associated with the development of reading skill in general (Manis, Seidenberg, & Doi, 1999; Morris et al., 1998; Scarborough, 1998; Wolf, Bally, & Morris, 1986; Wolf & Bowers, 1999) and potentially involved in the orchestration of processes involved in reading fluency (Wolf, Bowers, & Biddle, 2000). Rapid object-naming speed shares considerable independent variance with word identification (Bowers & Swanson, 1991). For the purposes of the present study, naming-speed deficits have also been shown to affect the rate with which individuals are able to read connected text (Bowers, 1993; Breznitz & Berman, 2003; Katzir, Shaul, Breznitz, & Wolf, 2004; Stage, Sheppard, Davidson, & Browning, 2001; Young & Bowers, 1995) and so may have particular implications for the development of reading fluency. Thus, in the development of reading, naming-speed problems result in the "slower access to lexical and sublexical information that may impede the development of fluency in reading" (Wolf et al., 2000, p. 401).

Orthographic processing

Orthographic processing can be defined as "the ability to represent the unique array of letters that defines a printed word, as well as general attributes of the writing system such as sequential dependencies, structural redundancies, and letter position frequencies" (Vellutino, Scanlon, & Tanzman, 1994, p. 314). Orthographic coding tasks are designed to assess children's knowledge of visual spelling patterns in a way that distinguishes them from the simple retrieval of letter sounds. Orthographic processing differences among children may account for additional variance in word-reading skill independent of that accounted for by phonological processing, such as nonword reading (Cunningham & Stanovich, 1990), but this is not always found (Manis, Custodio, & Szeszulski, 1993). Low-skilled readers may use this knowledge of visual spelling patterns to compensate for otherwise poor phonics skills as an alternative route to word identification (Pennington, Lefly, Van Orden, Bookman, & Smith, 1987; Stanovich & Siegel, 1994), but this point is also controversial (see Foorman, Francis, Fletcher, & Lynn, 1996).

In the current study, we assessed orthographic knowledge in two ways. First, we used the Experimental Spelling Task, developed by Olson, Kliegl, Davidson, and Foltz (1985), to assess children's orthographic processing. In this task, children are asked to decide which of two letter strings represents the correct spelling of a real word, such as *rane* versus *rain*. Because both strings obey the spelling–sound correspondences of English and both sound like a real English word when sounded out, children must retrieve orthographic information to discriminate between them. Variants of this basic task have been used by a number of researchers to measure orthographic knowledge (Manis et al., 1993; Olson et al., 1985; Stanovich & Siegel, 1994). Some researchers (Manis et al.; Vellutino et al., 1994) have taken issue with this means of assessing orthographic knowledge because it is possible for children to make these decisions on the basis of their prior experience reading the words, their spelling instruction, or the mere presence of more orthographic "neighbors" (i.e., *rain*, *vain*, *stain*, *grain*, *brain*) in their mental lexicon for one option over the other. Still, it is probably the most widely used task to assess orthographic knowledge, so we included it in our study.

Second, we used a variant of the Doublet Knowledge Test, developed by Cassar and Treiman (1997), to assess metacognitive knowledge of doublets. This task assesses children's understanding of the permissibility of consonants and vowels as doublets with regard to position (i.e., consonant doublets tend not to occur at the beginning of a word with a rare exceptions such as *llama*) and letter (i.e., *ii* and *uu* cannot appear as doublets within a word; some consonants such as *hh*, *jj*, and *vv* are rarely doubled). This task may also have some of the problems of Experimental Spelling Task (i.e., neighborhood effects) but seems to relate less ambiguously to true orthographic knowledge. In our variant of this task, participants were asked to decide as quickly as possible which of two items such as *baff* versus *bbaf* "looks more like a word should look." Cassar and Treiman found that this type of knowledge emerges in kindergarten and is fully accurate by second grade. In our variant of this task, we presented children with the doublet task and asked them to make the doublet decision as quickly as they could.

Text-reading fluency

Typically, assessments of text-reading fluency entail asking children to read aloud from connected text while an examiner records, at minimum, the number of correctly read words per minute. At issue is exactly what additional benefit text-reading fluency provides beyond those accrued from skilled, fluent word recognition.

There are a number of potential benefits that the fluent reading of connected text can provide over reading isolated words. Among them, fluent oral readers verbally segment word strings into larger syntactic groupings that may support comprehension (Dowhower, 1987; Kuhn & Stahl, 2003; Rasinski, 1990; Wolf & Katzir-Cohen, 2001; Young & Bowers, 1995). Text reading makes available context-facilitated word recognition as well (Stanovich, 1980). To the extent that these processes are carried out automatically, children can gain additional resource benefits from text reading that they can use for comprehension. The point of view represented by this model is supported by studies finding that text-reading fluency accounts for additional variance in reading comprehension beyond that accounted for by word-reading fluency alone (Jenkins et al., 2003). However, whether children in the early phases of learning to read are able to make use of the resource benefits ascribed to text-level reading is controversial (Kuhn & Stahl, 2003; Schwanenflugel et al., 2004; Shinn et al., 1992). In the current study, to assess text-reading fluency, children read passages from the Gray Oral Reading Test–Third Edition (GORT–3). This assessment is perhaps the most widely used standardized assessment of reading fluency (Marlow & Edwards, 1998).

The autonomy component

The Stroop task directly taxes the autonomy component in early reading by requiring participants to name a primary stimulus (color or picture) while ignoring a printed word (e.g., naming the picture of a desk with the word *cat* written on it, or naming the ink color “blue” for the word *red* printed in blue ink). In the picture–word version of the task, interference results when the reader cannot help but read the printed word while attempting to name the picture. Interference is determined by comparing the time it takes to complete the naming of the primary stimulus in a Stroop condition against the naming of a control having random letter strings or no conflicting print at all.

Several studies have linked the development of Stroop interference to the development of reading skill. Studies by Stanovich, Cunningham, and West (1981) and Guttentag and Haith (1978) found an increase in Stroop interference from the beginning to the end of the first-grade year. Ehri (1976) found that early and proficient readers showed larger interference, compared with late and less proficient readers. Similarly, Schadler and Thissen (1981) found that Stroop interference from words increased until children’s reading skills reached a fourth-grade reading level, after which interference declined (see also Golinkoff & Rosinski, 1976). In later elementary school, as children become better able to inhibit distractions, the level of interference is decreased (Pritchard & Neumann, 2004; Tipper, Bourque, Anderson, & Brehaut, 1989). Although the overall size of the Stroop effect may change, the effect itself does not disappear across the life span (Comalli, Wapner, & Werner, 1962).

Given that Stroop interference seems to develop along with reading skill, what exactly becomes processed autonomously as reading becomes automatic? Some early findings imply that not only are words processed autonomously but also letter units may be (Guttentag & Haith, 1978). Some researchers hypothesize that children begin the process of learning to read by using small units that map letters onto distinct phonemes, called grapheme–phoneme correspondence rules (or GPC; Coltheart & Leahy, 1992; Perry, Ziegler, & Coltheart, 2002). A grapheme is a single letter (e.g., *t*-) or letter cluster (*sh*-) that gets mapped on to a single phoneme or sound. According to this view, children initially read using GPC units, making use of larger letter–sound mappings only later when they have built up a large sight-word vocabulary. An alternative view is that children begin to read using larger sublexical units that engage either word analogies or rime units for word decoding (Goswami, Ziegler, Dalton, & Schneider, 2003). Rime units are the part of a single-syllable word used in a rhyme (e.g., c*at*, th*at*; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). Stroop interference might be expected to be different depending on the size of the units used. In the current study, we presented children with pictures containing pronounceable nonwords, words containing high-probability GPC units, words containing high-probability rime units, and words containing both.

Reading comprehension

Finally, to determine whether there were benefits from fluent text reading, we needed to have an assessment of reading comprehension. For this, we assessed children's comprehension using the reading comprehension subtest of the Wechsler Individual Achievement Test (WIAT). We selected this test because of its good psychometric properties and because it seemed to cover a very basic ingredient of reading comprehension, that is, that children should be able to answer comprehension questions (both simple and more inferential) following the reading of a text. Being able to answer questions regarding literal and inferential information following the reading of a text has been demonstrated to be a key discriminator between skilled and less skilled comprehenders (Cain, Oakhill, & Bryant, 2004; Casteel, 1993; Long, Oppy, & Seely, 1997).

To sum, the purposes of the present study were two. The first was to develop an empirically based model of the development of fluent and automatic reading in the early elementary school years. In particular, we were interested in determining how lexical and sublexical processes (word reading, nonword reading, orthographic processing, and rapid object naming) and reading autonomy (Stroop) operate together to produce the development of fluent text reading and good comprehension of text. The second was to determine whether fluent text reading provided additional benefits for reading comprehension beyond those skills involved in word-reading fluency alone. First-, second-, and third-grade children were asked to complete a series of reading tasks targeting reading fluency, reading autonomy, and reading comprehension. Performance on each task was examined for developmental change across grade levels. Then, structural equation modeling was carried out to evaluate how these factors operate together to produce fluent text reading and good comprehension.

Method

Participants

Participants were 99 first-grade (mean age = 7 years, 1 month; range = 5 years, 7 months to 8 years, 2 months), 79 second-grade (mean age = 8 years, 1 month; range = 7 years, 1 month to 9 years, 1 month), and 71 third-grade (mean age = 9 years, 0 months; range = 7 years, 11 months to 10 years, 4 months) children attending four public schools located in communities in urban northeast Georgia or suburban central New Jersey, USA. Approximately 41% were African American, 27% were European American, 23% were Hispanic American, 5% were Asian American, and 4% were other or unknown; 53% were female and 47% were male. All children in general education classrooms participated with the exception of those currently receiving English-as-a-second-language instruction or those in self-contained classrooms for special education. Children were not excluded on the basis of reading disability. The children attended schools in which approximately 76% received free and reduced-cost lunch.

Stimuli and procedures

Children were given a mixture of standardized and experimenter-constructed reading measures during the spring term (seventh and eighth month) of the school year. The order of the standardized measures was counterbalanced with the experimental reading measures, such that half the children received the standardized measures first and the other half received the experimenter-constructed reading measures first.

Standardized reading measures—The standardized reading measures were the following:

a. *Test of Word Reading Efficiency (TOWRE) Sight Word Efficiency and Phonemic Decoding Efficiency subtests, Form A (1999).* The Sight Word Efficiency subtest asks children to name all of the words they can from a list of words in 45 seconds. The Phonemic Decoding Efficiency subtest requires children to pronounce all of the phonetically regular nonwords they can in a list in 45 seconds. Concurrent validity estimates reported in the test manual have a median of .91 in grades 1 through 3. Alternate form reliabilities have a median of .97 in grades 1 through 3.

b. *Wechsler Individual Achievement Test (WIAT; 1992) Reading Comprehension subtest.* The WIAT Reading Comprehension subtest consists of a number of passages that increase in complexity. Moreover, passage questions move from assessing literal to inferential comprehension, consistent with research on the development of reading comprehension (Cain et al., 2004; Hannon & Daneman, 2001). In this assessment, a comprehension question is asked when the child completes the reading of a passage, and the child answers the question aloud in his or her own words. The WIAT is discontinued once the child misses four sequential questions. Validity estimates of this subtest reported in the test manual have a median of .74 and a reliability of .91 in grades 1 through 3.

c. *Gray Oral Reading Test–Third Edition (GORT–3, 1992).* The GORT–3 is an assessment of proficiency in the oral reading of text. The test is designed to evaluate text-reading fluency and consists of a series of passages that increase in difficulty. The test has several subscales, but we focused on reading rate for the purpose of this study. Accuracy and rate are tabulated for each passage until the reading becomes too slow and inaccurate and meets the discontinue rule established by standardization. The reading of each passage received a rating for reading rate assigned by the test makers that is based on standardized data regarding how long it took children to read each passage. These ratings are then summed to form a cumulative reading rate score. This score served as our indicator of text-reading fluency. Validity estimates for the passage-reading rate scores presented in the test manual range from .34 to .82, with a median validity of .65. The test has been independently validated for a diverse population of children (Craig, Thompson, Washington, & Potter, 2004). Reliability estimates for the assessment of passage-reading rate range from .82 to .92, with a median reliability of .90.

d. *Comprehensive Test of Phonological Processing Rapid Object Naming subtest (CTOPP–RON; 1999).* The Rapid Object Naming subtest of the CTOPP–RON measures the time children take to name six different objects alternately presented randomly over 72 objects. The test is said to be a measure of the efficient retrieval of visual and phonological information from long-term memory. The manual reports reliability estimates for children in the current age range from .73 to .93 on this subtest, with a median of .77, and validity estimates against reading measures ranging from .30 to .54, with a median of .44.

Experimenter-constructed reading tasks—The experimenter-constructed reading tasks were designed to provide a more fine-grained assessment on aspects of early reading not covered by the standardized measures of reading, particularly those related to the automaticity aspects of reading. These tasks were all presented on a Dell Inspiron 8000 lap-top computer using E-Prime experiment software (from Psychology Software Tools, Version 5.0; Schneider, Eschman, & Zuccolotto, 2001), which was connected to a serial response box (Psychology Software Tools, Model # 200A). Children were tested in a quiet location in their school. Each participant was seated in front of the computer screen and asked to hold the microphone/ response box while the experimenter read a set of instructions aloud from the computer screen for each task. For tasks using the microphone, the accuracy of participants' responses was recorded directly by the experimenter as incorrect if the item was not appropriately named. It was scored as a mechanical error if the microphone was not triggered by the child's voice or if it was triggered by an extraneous noise unrelated to the experiment (e.g., background noise, cough, sneeze, pre-articulation response such as "tsk").

(a) Stroop task: The purpose of the Stroop task was to provide an indicator of the degree to which word reading was autonomous in the children. The Stroop task presented digitized line drawings of objects provided in Cycowicz, Friedman, Rothstein, and Snodgrass (1997), each with a single conflicting word written in lowercase letters over the middle of it (e.g., a picture of a cat with the word *rock* superimposed on it). There were six conditions varying in the grapheme–phoneme (GPC; Berndt, Reggia, & Mitchum, 1987) and rime-unit relationships represented by the words superimposed on the pictures: control pictures without words, control pictures with random letter strings, pictures having words with highly predictable GPC and highly consistent rime units, pictures having words with highly predictable GPC units but with inconsistent rime units, pictures having words with low-predictability GPC units but with highly consistent rime units, and pictures having nonwords that could be pronounced with highly consistent rime and predictable GPC units. Words were chosen so that they did not differ in normative word frequency across conditions according to Zeno, Ivens, Millard, and Duvvuri (1995), $p > .10$. Items were presented in random order. There were 6 practice and 60 experimental stimuli.

For each trial, a fixation point appeared when the experimenter pressed a button on the serial response box, followed a second later by the picture stimulus. Participants then named each picture stimulus aloud into the microphone. The accuracy of the participant's response was recorded by the experimenter. Spearman-Brown coefficient analysis indicated a split-half reliability on this task of .89 for reaction times and .79 for accuracy rates.

(b) Single-word–naming task: The purpose of the single-word–naming task was to provide an indicator of word-reading speed for a variety of word types while controlling for part of speech. With the use of normative ratings collected by Schwanenflugel and Noyes (1996) from third-grade children, the stimuli selected were a mixed set of nouns ranging in imageability ($M = 4.5$, $SD = 1.7$, range 1.4–6.8 where 1 means "hard to think of a picture for the word" and 7 means "easy to think of a picture for the word"), rated context availability ($M = 5.5$, $SD = .7$, range 4.1–6.6 where 1 means "hard to think of a sentence for the word" and 7 means "easy to think of a sentence for the word"), frequency ($M = 63$, $SD = 44$, range 1 to 171 in the third-grade corpus of Carroll, Davies, & Richman, 1971), and orthographic regularity based on Venezky (1999). Thus, this task included words ranging in semantic and orthographic difficulty.

Participants were presented with words one at a time. Each child was asked to "say the word as quickly as you can when the word comes up on the screen." For each trial, a fixation point appeared for one second in the middle of the computer screen, and this was followed by the presentation of the word. When the word was named, it disappeared from the screen, reaction time (RT) was recorded automatically, and the accuracy of the response was keyed in by the experimenter. Pilot testing had indicated that the task was difficult for some children; therefore, we decided to include a ceiling rule so that the children with less reading skill would not be unduly frustrated by the task. If the child made more than six errors in a row, the task was discontinued. To ensure that there were observations from each condition, items were presented in a fixed order randomized within blocks of four trials, randomly ordering condition within each block. There were 8 practice and 48 experimental items. Spearman-Brown coefficient analysis indicated a split-half reliability on this task of .80 for reaction times and .97 for accuracy rates.

(c) Doublet knowledge task: The purpose of the doublet knowledge task was to provide an indicator of the degree to which knowledge of spelling patterns could be accessed automatically with increasing reading fluency. The task was modeled after the Cassar and Treiman (1997) Doublet Knowledge Task with the exception that the current task was timed to better determine automaticity of this knowledge. Children were given pairs of nonword letter strings having consonant or vowel doublets (e.g., *holl* versus *hhol*, *stee* versus *staa*) and asked to "pick the item that looks more like a word should look." For each trial, a fixation point appeared for one second and was followed by the letter-string pair. If the item on the left was chosen, then the child pressed the left button on the response box; if the one on the right was chosen, then the child pressed the button on the right. The items were selected from ones used by Cassar and Treiman such that there were 8 practice and 22 experimental items. RT and accuracy were recorded automatically. Spearman-Brown coefficient analysis indicated a split-half reliability on this task of .91 for reaction times and .73 for accuracy rates.

(d) Experimental spelling task: The purpose of the experimental spelling task was to provide an additional assessment of the automaticity of accessing orthographic knowledge. The experimental spelling task was adopted from one developed by Olson et al. (1985; Olson et al., 1991). In this task children were provided with item pairs, one of which was a correctly spelled version of a word (e.g., *rain*) and the other of which was a phonologically identical but incorrectly spelled version of the word (e.g., *rane*). The children were told, "Two groups of letters will appear on the computer screen. Both sound like a real word, but only one is a real word. Pick the one that you think is the real word by pushing the button that is on the same side as the real word.... Do this as quickly as you can without making mistakes." There were 8 practice and 22 experimental trials. Half of the correct spellings appeared on the right and the other half on the left. RT and accuracy was recorded automatically. Spearman-Brown coefficient analysis indicated a split-half reliability on this task of .87 for reaction times and .69 for accuracy rates.

Results

The findings were analyzed in two phases. First, we carried out simple cross-sectional analyses to assess developmental change in the computerized tasks and standardized tests. Then, with knowledge of how children progressed on the individual tasks, we carried out tests of theory using structural equation modeling of the relationships among the tasks.

Analyses of developmental change on tasks and tests

Analyses of standardized measures—We began by carrying out basic analyses of variance of the standardized test data (TOWRE–SWE and PDE, WIAT–RC, CTOPP–RON, and GORT–3) using Grade as a between-subjects factor, focusing on raw scores for the tests. The raw score means and standard deviations for these subtests and their corresponding standard scores can be found in Table 1.

For each assessment, there was improved performance throughout the early elementary school grades. For the TOWRE, there was a significant main effect of Grade on both the SWE subtest, F (2, 246) = 83.45, and the PDE subtest, F (2, 246) = 57.197, both $p < .001$. Bonferroni post-hoc follow-up tests indicated that third graders read significantly more sight words and nonwords than second graders, who, in turn, read more words and non-words than first graders, all $p < .001$. For the WIAT–RC, there was also significant main effect of Grade, F (2, 246) = 80.21, $p < .001$. Again, post-hoc tests indicated that older children answered more passage questions correctly than younger children, all $p < .001$. Finally, for the CTOPP–RON, there was a significant main effect of Grade, F (2, 246) = 22.12, $p < .001$. Post-hoc tests indicated that first graders took longer to name all the pictures than second graders, who took longer than third graders, all $p < .05$. For the GORT–3, there was a significant main effect of grade on reading rate sum scores, F (2, 246) = 75.74, $p < .001$. Post-hoc tests indicated that first graders read the texts at a slower rate than second graders, who in turn read more slowly than third graders, all $p < .001$.

Computerized reaction time (RT) tasks—For all computerized RT tasks, to identify RTs and error rates that qualified for further analysis, the following steps were conducted on the raw data. First, all RTs for incorrect responses were eliminated from the data and their responses counted as errors. Then, all mechanical errors (as defined above) and their corresponding RTs were eliminated from the data. Next, mean RT and standard deviations were calculated for every participant in the study. Then, RT trials greater than 2 standard deviations above each participant's mean and less than 200 ms were removed from the data as outliers. Finally, for each participant, means were recalculated for RT and error rates for each condition. These means formed the basic data for subsequent analyses (see Table 2).

(a) Stroop task: Because error rates approached floor for all conditions and grades, we focused on picture-naming time means as the major dependent variable of interest. A 3 Grade × 6 Stroop Condition ANOVA was carried out with Grade as a between-subjects factor and levels of Stroop (picture control, letter-string control, nonword labels, Low-GPC + High-Rime Labels, High-GPC + High-Rime Labels, High-GPC + Low-Rime Labels) as a within-subjects factor. There was a significant main effect of Grade, $F(2, 246) = 6.89$, suggesting that third-grade children named the pictures more quickly in general than first- and second-grade children. It is important to note that there was a significant main effect of Stroop, $F(5, 1230) = 23.57$, $p < .001$, and a non-significant interaction between Stroop and Grade, $F(10, 1230) = 1.39$, $p = .178$, suggesting similar patterns of interference across grade. Further, we dropped the picture control condition so that the influence of sublexical units beyond the mere letters could be examined directly. This analysis, too, found a significant main effect of Stroop Condition, F (4, 984) = 15.01, $p = .003$, and Grade, $F(2, 246) = 5.85$, $p = .003$, and no interaction between Grade and Stroop Condition, $F(8, 984) = 1.12$, $p = .348$, suggesting similar patterns of interference across grade. Follow-up Bonferroni tests, adjusting alpha by the number of contrasts (.05/4 = .013), indicated that compared to the letter-string controls, non-words, High-GPC + Low-Rime words, and High-GPC + High-Rime words interfered significantly with picture-naming time, but not Low-GPC + High-Rime words. This suggests that the operative sublexical reading unit in Stroop effects in early elementary school is the presence of a highly predictable GPC unit.

The presence of Stroop interference is generally recognized as the operation of autonomous processing of words, the degree of which varies across participants. Thus, for the remainder of the analyses, we calculated Stroop interference directly for each participant by subtracting each one's mean RT for letter strings from their RT from nonwords, High-GPC + Low-Rime, and High-GPC + High-Rime words. These interference scores served as indexes of Stroop interference for the assessment of the role of autonomous reading in subsequent analyses.

(b) Single-word–naming task: The findings for the single-word–naming task indicated that younger children found decoding this set of words slow and difficult. A perusal of error rate means indicated that errors were not at floor, so a separate analysis of them was conducted as well. A one-way ANOVA indicated a significant main effect of Grade for both naming time, $F(2, 246) = 55.25$, and error rates, $F(2, 246) = 93.28$, both $p < .001$, such that word decoding became quicker and more accurate for this set of words throughout the early elementary school years. So that text-reading and word-reading skill could be more directly compared, RTs were converted into words/minute and error rates were converted into accuracy rates. Thus, for subsequent analyses, words per minute (WPM) and accuracy rates served as indicators of single-word–naming skill.

(c) Doublet Knowledge Task: For this task, we were interested in the speed and accuracy with which children could make decisions regarding which doublet nonword "looked more like a word should look." ANOVA using Doublet Type (vowel or consonant) as a within-subjects factor and Grade as a between-subjects factor indicated a significant main effect of Doublet Type, $F(1, 246) = 36.61$, such that it took children longer to decide that vowel doublets looked more like a word should look than consonant doublets. There was a nonsignificant Doublet Type × Grade interaction, $F(2, 246) = 2.91$, $p = .056$, for RTs. There was a main effect of Grade for decisions regarding consonants for both RT, Grade, $F(2, 246) = 3.34$, and error rates, Grade, $F(2, 246) = 47.11$, both $p < .05$, such that third graders took less time and made these decisions more accurately than first graders. There was a main effect of Grade for decisions regarding vowels for accuracies, Grade, $F(2, 246) = 3.14$, $p = .045$, but not for RT, Grade $F(2, 246) = 2.21$, $p = .112$. Condition means on RTs and accuracies served as indicators for the development of doublet knowledge in younger and older children in subsequent analyses.

Experimental spelling task—For this task, we were interested in whether children's knowledge of correct spellings could be accessed quickly with age. An analysis of variance comparing mean RTs and error rates for this task as a function of Grade indicated a significant main effect of Grade on RTs, $F(2, 246) = 9.72$, and error rates, $F(2, 246) = 98.64$, both $p < .001$. Third graders accessed this knowledge more quickly than first graders and more accurately than children in the earlier grades ($p < .01$). Second graders made these decisions more accurately than first graders ($p < .001$). For subsequent analyses, mean RTs and error rates served as indicators of the development of spelling knowledge in these children.

In sum, between first and third grade in early elementary school, we found evidence of development on all tasks potentially related to fluent reading and automatic reading in young children. Compared with younger children, third-grade children pronounced sight words and nonwords at a faster rate on the TOWRE, read text passages at a faster rate on the GORT–3, and read single words more quickly and accurately. Older children made doublet decisions regarding consonants more quickly and accurately and regarding vowels more accurately. They determined correct spellings of words more quickly and accurately. They also named pictures more quickly on the CTOPP–RON. Moreover, third graders could answer more comprehension questions correctly after reading passages than first graders on the WIAT–RC. The sole exception to this was the pattern of developmental change on the Stroop task, which found similar patterns of interference with Stroop interference as a function of grade, although third graders showed less interference overall.

Structural equation modeling

Structural equation modeling in LISREL was conducted to examine the relations between the various measures potentially related to the development of fluency during the early elementary school years. The structural equation modeling was carried out in two phases to evaluate particular models: (a) a measurement modeling phase to identify whether the observed variables fit the latent factor structure proposed by theory, and (b) a structural modeling phase to test the explanatory relationships indicated between predictor and outcome variables by particular models (Jöreskog & Sörbom, 1996).

Measurement modeling phase—In this phase, relations among variables thought to be potentially relevant to reading fluency according to variants of automaticity theory were determined using LISREL being as inclusive as possible with regard to the variables that might be considered to be relevant. In the measurement modeling phase, we included all tasks. We used raw scores from the standardized assessments rather than age-based standard scores because the correlations between the raw and age-based standard scores correlated > .90, but the raw scores allowed us to evaluate the grade effects directly in terms of potential changing relationships between variables as a function of age. In this initial phase, text-reading fluency and reading comprehension were considered outcome variables. We defined our interest in reading fluency to be those aspects of reading that relate to better and worse comprehension.

The correlation matrix for each grade level was used separately as input to LISREL 8.30 (Jöreskog & Sörbom, 1996, 1999) and is the focus of the analysis. We considered that autonomous reading would consist of a latent factor comprising the mean difference in RT between letter controls and the three conditions in which the interfering stimuli contained a highly predictable GPC unit (High-GPC + Low-Rime, High-GPC + High-Rime, and Nonwords). We assumed a latent factor of word-reading fluency comprising the number of sight words read for the TOWRE sight-word efficiency, the mean WPM calculated for the single-word–naming task, and the mean accuracies for that task. We also included in this latent word-fluency factor the number of nonwords read in TOWRE phonemic-decoding efficiency and the number of seconds to complete the CTOPP–RON task. As noted earlier, nonword reading is generally considered a key indicator of a child's ability to read unknown words. Further, rapid autonomous naming is thought to share many of the skills underlying word-identification speed, such as quickly perceiving stimuli, retrieving symbolic information from long-term memory, and phonological production (Wolf et al., 2000). Thus, it seemed appropriate to have this variable serve as an indicator of word-reading fluency. Finally, we also included in this latent factor the RT and accuracy means for all orthographic knowledge task conditions because orthographic knowledge is thought to be important to the fluent identification of words (Cassar & Treiman, 1997; Olson et al., 1985). Both of these factors (word-reading fluency and autonomous reading) served as predictors of the GORT–3 and the WIAT.

Measurement modeling was carried out for each grade separately. We used several indicators for lack of model fit recommended by Kline (2005) and Schumacker and Lomax (1996). It is traditional to present the model χ^2 fit index, which tests "the difference in fit between a given overidentified model and a just-identified version of it" (Kline, 2005, p. 136). A nonsignificant $p > .05$ is an acceptable fit. This is generally not the preferred fit index for various reasons and needs to be supplemented with other indicators of fit. We included the root mean square error of approximation (RMSEA) as an assessment of simple model misspecification where RMSEA < .08 indicates an acceptable and < .06 a good fit. We also provide the goodness of fit index (GFI), an absolute fit index that estimates proportion of variability in the sample matrix explained by the predicted covariance matrix, and the Tucker–Lewis index (TLI), an incremental fit index sensitive to misspecified factor loadings that is parsimony adjusted. For these, fits > .90 are considered acceptable and > .95 a good fit.

Overall, measurement model results indicated an unacceptable fit for first grade, χ^2 (100) = 256.71, $p < .001$; RMSEA = .11; GFI = .77; TLI = .80. The same was true for second grade, χ^2 (100) = 326.70, $p < .001$; RMSEA = .17; GFI = .66; TLI = .61; and for third grade, χ^2 (100) = 276.69, $p < .001$; RMSEA = .14; GFI = .70; TLI = .57. However, a perusal of fitted standardized residuals indicated the source of the ill fit. We found that large residuals (> 2.58) were consistently associated with the orthographic tasks at each grade level (5 of 5 for first grade, 7 of 10 for second grade, and 7 of 9 for third grade). Although we made several other attempts to include these variables in the model, every other attempt produced the same result. Thus, the results of this measurement modeling indicated that it was best simply to exclude these orthographic variables from further model testing. When these variables were excluded, the models obtained reached an acceptable fit to data for first, χ^2 (33) = 40.44, p = .170; RMSEA = .034; GFI = .93; TLI = .99; and third grade, χ^2 (33) = 51.30, p = .022; RMSE = .059; GFI = .90; TLI = .92, although the fit for the second grade did not quite meet established levels, χ^2 (33) = 62.18, p = .002; RMSEA = .10; GFI = .87; TLI = .91. However, in the interest of parsimony, we deemed it better to proceed by testing the same models for all grades.

Tests of the hypothesized structural models—Structural equation modeling was used to examine the role of word-reading fluency, autonomous word reading, and text-reading fluency in benefiting reading comprehension skills. We tested two theoretical versions of the way that these skills related to one another: (a) text reading as mediator model and (b) simple reading fluency model.

(a) Text reading as mediator model: In this model, the autonomous reading interference and word-reading–fluency aspects of reading were viewed as predictors of text-reading fluency and reading comprehension. As noted above, the latent factor reflecting word-reading fluency was indicated by both subtests of the TOWRE, single-word–naming accuracy and WPM from the single-word–naming computerized task, and the CTOPP–RON. This latent factor served both as a predictor of text-reading fluency and reading comprehension. However, in this model, text-reading fluency was considered a mediator of the relation between word-reading fluency and reading comprehension. As noted earlier, the mediating quality of text reading is said to emerge from the special features of text reading not available from fluent word reading alone. A second latent factor consisting of the three interference scores from the Stroop task served as an indicator of the autonomy of lexical processing. To the extent that the presence of autonomous word reading is a component of automaticity, it should be related to word-reading fluency, text-reading fluency, and, perhaps, reading comprehension.

The proposed theoretical model was submitted to LISREL. The resulting analysis provided an acceptable fit according to most of the fit indexes for the first-grade data, χ^2 (31) = 37.54, p = .190; RMSEA = .03; GFI = .93; TLI = .99; and third-grade data, χ^2(31) = 50.10, p = .016; RMSEA = .067; GFI = .90; TLI = .91, but a less acceptable fit for the second-grade data, χ^2(31) = 61.99, $p < .001$; RMSEA = .11; GFI = .87; TLI = .91. However, given the similarity between first- and third-grade data, it seemed more parsimonious to examine the same model for second grade. The model is presented in Figures 2a, 2b, and 2c, and exact t-levels for standardized estimates for paths in Table 3; all factor loadings for each grade were significant at the $p < .05$ level. Solid arrows represent significant paths and dashed lines represent nonsignificant paths.

An examination of the paths leading to text-reading fluency as an outcome measure indicated a significant direct path between the latent word-reading–fluency factor and text-reading fluency for all grades ($p < .05$). In contrast, the direct path between the autonomous reading interference latent factor and text-reading fluency was not significant at any grade (all $p > .10$). Moreover, the correlation between autonomous word reading and the word-reading–fluency latent factors was small and non-significant at first (r = .06), second, (r = .06) and third grades (r = .06). Taken together, the structural equations for this model accounted for 61% of the variance in text-reading fluency in first grade, 74% of the variance in second grade, and 73% of the variance in third grade.

Examination of the direct paths leading to reading comprehension as an outcome measure indicated the importance of word-reading fluency in reading comprehension. The direct path between the word-reading fluency factor and reading comprehension was significant at all grades ($p < .05$). Moreover, the path between autonomous reading interference and reading comprehension was significant for first graders ($p < .05$) but nonsignificant for the other grades, suggesting this aspect of automaticity is not important once basic automaticity is established. It is important, however, that this theoretical model predicts that the indirect path from the text-reading fluency to reading comprehension outcomes should be significant. However, at all grades, this relationship was nonsignificant (all $p > .10$). Thus, the anticipated additional variance from text-reading fluency for comprehension beyond that predicted from word-reading fluency predicted by this model simply was not confirmed by this data. Thus, if there are important contributors of fluent text reading on comprehension, it is unclear what they are for young readers. Taken together, the amount of variance accounted for in reading comprehension by the structural equations in this model decreased from first (76%), to second (45%), to third grade (38%).

Simple reading fluency model—This model would state that reading fluency is a rather general trait of skilled readers and that, at least for young readers building fluency, it is a general trait that predicts comprehension. This model does not distinguish text-reading from word-reading fluency skill and represents text-reading fluency as merely another indicator of reading fluency. Thus, in this model, children's scores on both subtests of the TOWRE, mean single-word–reading time and accuracy on the single-word–naming task, CTOPP–RON, and scores on the GORT–3 all served as indicators of a reading fluency factor.

Similar to the previous model, it was predicted that autonomous reading interference should be correlated with the development of reading fluency, particularly as reading automaticity emerges. This view would also seem to state that the emergence of all aspects of automaticity, that is, fluent reading and autonomous reading, should be related to improved comprehension. Thus, as before, children's interference scores for the Stroop task served as indicators of the autonomous reading factor.

We should note that this and the previous model are identical in most respects with the exception that the parameters linking text-reading fluency to reading comprehension and automatic reading to text-reading fluency are now eliminated. With two fewer parameters included, this model is simpler and more parsimonious than the previous one.

The proposed theoretical model provided a reasonable fit according to most of the fit indexes for the first-grade data, $\chi^2(33) = 40.44$, $p = .170$; RM-SEA = .034; GFI = .93; TLI = .99; and third-grade data, $\chi^2(33) = 51.30$, $p = .022$; RMSEA = .059; GFI = .90; TLI = .92, but a less acceptable fit for the second-grade data $\chi^2(33) = 62.18$, $p = .002$; RMSEA = .10; GFI = .87; TLI = .91. However, as before, given the similarity between first- and third-grade data, it seemed more parsimonious to examine the same model for second grade. All factor loadings were significant at the $p < .05$ level for all relevant variables at each grade level. Standardized path weights and *t*-values can be found in Table 4. All factor loadings for each grade were significant at the $p < .05$ level.

As noted in Table 4 and Figures 3a, 3b, and 3c, for all grades there was a significant path between reading fluency and reading comprehension ($p < .05$) indicating that, as children become fluent readers, they also comprehend what they read better. As before, autonomous reading interference was uncorrelated with reading fluency at first grade, $r = .03$, and second grade $r = .07$, $p > .10$, but, by third grade, children who continued to show large amounts of autonomous reading interference tended to display significantly worse reading fluency ($r = -.24$, $t(31) = 2.30$, $p < .05$). Autonomous reading interference was a significant predictor of reading comprehension in first grade, $t(31) = 2.33$, $p < .05$, but not in second or third grades, both $p > .10$. Thus, what utility autonomous reading aspect of reading has for reading comprehension disappears once basic reading fluency is established. Moreover, continued difficulty recovering from Stroop interference seems to be associated with difficulties with establishing reading fluency.

It is important to note, however, that an examination of the variance accounted for in reading comprehension by this structural equation model pointed to the declining influence of general reading fluency and autonomous reading on reading comprehension with age. Taken together, the amount of variance accounted for in reading comprehension by the structural equations in this model decreased from 75% in first grade, to 45% in second grade, and 39% in third grade. This supports the common observation that, as children become older and their texts become more difficult, we need to look beyond reading fluency to other skills and knowledge (i.e., vocabulary knowledge, world knowledge, inferencing ability) to account for reading comprehension ability.

Discussion

Despite the popularity of automaticity views regarding the development of reading fluency over the past 30 years, there have been remarkably few investigations examining the interrelationships between the proposed features of automaticity and the development of reading. Most typically, investigations have focused on two or three elements of this view simultaneously (e.g., Fuchs et al., 2001; Jenkins et al., 2003; Stanovich et al., 1981; Young & Bowers, 1995). Although it is impossible to capture in a single study all of the elements that might be relevant to the developing role of reading fluency in reading comprehension, the current study made a comprehensive attempt to capture most of the important factors on which automaticity views have focused. This study provides some clarification regarding how the elements of automaticity, word-reading fluency, text-reading fluency, and autonomous reading, operate together to produce good comprehension in the early elementary school years.

Our study examined the suitability of two versions of the automaticity view of reading. For the *text reading as mediator model*, fluent readers are said to be better able to utilize text fluency-related skills such as sentence-parsing skills (Schreiber, 1980) and the contextual activation of word meanings (Jenkins et al., 2003) to aid in the comprehension of text once their basic word reading is automatic enough to free up cognitive resources. The current study found no evidence for this mediating role of text-reading fluency. Instead, a simpler model treating text-reading fluency as merely another indicator of reading fluency skill provided a more appropriate account for our data. Thus, our data supported a *simple reading fluency model* for the early elementary school years.

This lack of a mediating effect between word-reading fluency and reading comprehension skill for text-reading fluency is surprising given recent findings regarding the contribution of text fluency to reading comprehension (Cain et al., 2004; Jenkins et al., 2003). For example, Jenkins et al. found that text-reading fluency was the primary factor in predicting reading comprehension.

What can account for the differences between these results and ours? We think this is where a developmental account of reading skill must take place. In virtually all the studies finding additional benefits from text reading, children have tended to be older and more fundamentally fluent than the children in the current study (Cain et al., 2004; Jenkins et al., 2003; Nation & Snowling, 1998). Even the skilled older readers in our study were unable to use the additional cognitive resources gained by fluent text reading to comprehend text. We know that fluent early elementary school children use syntactically appropriate prosody and expression when they read (Clay & Imlach, 1971; Zutell & Rasinski, 1991). However, some research suggests that the full understanding of syntax is still under development during the age ranges we tested here (Traxler, 2002; Willows & Ryan, 1986). It may be that the imposition of syntactically appropriate phrasing during the reading of text is not yet automatic enough to free up resources that can be used for comprehension. Further, it is unclear whether children can orchestrate automatic word and text reading in the pursuit of better reading comprehension at this stage of reading (Wolf & Katzir-Cohen, 2001).

Comprehension is always an interaction between skills of the reader and characteristics of the text. Consequently, it may be that children at this age are, indeed, able to use their cognitive resources gained from fluent text reading to assist in comprehension, but that early texts are simply not complex enough to demand these additional resources. As children get older, the demands of the texts themselves will become greater (Hiebert & Fisher, 2005). It may be that it is not until later that children can use these additional benefits from text-level fluency to improve their comprehension.

Thus, we conclude that, in the early stages of the development of reading fluency, children use their word-reading skills to comprehend the meanings of text, so the *simple reading fluency model* applies. In fact, a study by Shinn et al. (1992) suggests that this may be the case. They found that, for third graders, a unitary factor placing word- and text-level skills and comprehension skills on the same factor provided the best fit for third graders, but by fifth grade, text-level skills had more closely aligned themselves with reading comprehension and required a multifaceted view of reading skill.

On a cautionary note, however, the sample sizes that we had available to us were somewhat smaller than ideal. One rule of thumb for deciding sample sizes in SEM is approximately 100 cases per sample. Another is 5 to 10 observations per parameter. Our sample sizes are clearly toward the lower end of these guidelines. Consequently, parameters that we did not find to be significantly significant or ill-fitting should not be dismissed completely out of hand. It was the case, however, that the critical parameter that distinguished the two models (i.e., the link between text-reading rate and reading comprehension) varied little across grade level, was very close to zero, and was in the opposite direction in two of the three cases. We think it unlikely that the addition of 20 or more participants at each grade level would have altered our findings dramatically.

One disappointment was our inability to discern the role of orthographic knowledge in the development of reading fluency and reading comprehension. We had supposed that including timed orthographic tasks might provide us with evidence that orthographic knowledge is related to other reading fluency skills. In support of this, we did find that older children were quicker and more accurate in making doublet decisions and in discerning the correct spelling of a word. This is generally consistent with other orthographic research using timed tasks (Gayan & Olson, 2003; Olson et al., 1985). However, as a group, the variables from our orthographic tasks were a source of ill-fit in the measurement modeling. Measurement issues are not uncommon in research on orthographic processing. For example, Cunningham, Perry, and Stanovich (2001) found low reliability on several commonly used orthographic tasks, making interpreting their role in reading difficult. Findings on one orthographic task may not always generalize to very similar sorts of tasks (Stanovich & Siegel, 1994). The dimensionality or covariance among orthographic tasks may also change during the early elementary school years (Gayan & Olson; Notenboom & Reitsma, 2003). Finally, the actual use of orthographic information during reading may be partly attributable to task features that encourage or discourage the use of orthographic information (Martensen, Dijkstra, & Maris, 2005). Thus, the mere validity of orthographic tasks such as the ones we used can be questioned (see also Vellutino et al., 1994). Consequently, we hesitate to say that orthographic information is *not* useful for fluent

NIH-PA Author Manuscript

NIH-PA Author Manuscript

reading and merely conclude that our current data do not provide us with a theoretical basis for deciding how it might be used.

A second potential issue for our conclusions is that our treatment of reading comprehension may not have been as complex as it needed to be in contrast with our treatment of reading fluency. It may have been that, if we had used a different measure of reading comprehension, we might have found a different role for text-reading fluency on comprehension. While this is possible, we believe our measure of reading comprehension included key features important in the early elementary school years. We selected the assessment that we did after evaluating a number of other standardized assessments of early reading comprehension. The WIAT Reading Comprehension test was chosen for its consistency with what many teachers consider a key indicator of reading comprehension (i.e., the child can answer questions about the text; Byers, 1998; Richardson, Anders, Tidwell, & Lloyd, 1991), while avoiding the presence of false test floors that include items outside of the domain of reading comprehension (e.g., single-word decoding only, letter recognition) that are sometimes found in standardized reading comprehension tests for early elementary school. Other research on this test suggested it correlates well with other tests of reading comprehension and that teachers view it as valid (Byers; Foegen, Espin, Allinder, & Markell, 2001; Smith & Smith, 1998). Moreover, the test we used assessed both literal and inferential comprehension, all important to reading comprehension (Cain et al., 2004). Still, it would be important to replicate these findings by using a broader range of comprehension measures than was present in the current study. Moreover, future research should consider the issue of the complexity of the to-be-comprehended texts in considering the relationship between text-reading fluency and comprehension.

A third potential issue for our conclusions is that our measure of text-reading fluency may not have been as complex as it needed to be because we relied heavily on a single text-reading measure, the GORT–3. However, reading fluency is typically defined as comprising text reading that is quick, automatic, and expressive. The current study lacked a measure of expressiveness during the reading of text. An earlier study by Schwanenflugel et al. (2004) had found that one measure of expressiveness, prosody, did not account for significant variance beyond word-reading fluency in predicting reading comprehension in second- and third-grade children. On the other hand, it is possible that other measures of expressiveness, such as those described by Cowie, Douglas-Cowie, and Wichmann (2002), may have accounted for significant variance above and beyond that accounted for by the GORT–3. The GORT–3 stresses rate and accuracy and not expressiveness. Expressiveness may be more relevant to the syntactic processing features of text fluency than rate and accuracy are. Future research should include a broader range of measures directly targeted at this expressiveness feature of text-reading fluency.

With these limitations in mind, having found support for a simple reading fluency model for early elementary school children, we outline our findings with regard to the development of each of the components of automaticity. We describe how these elements might fit together to produce good comprehension.

A number of descriptions of reading skill note the foundational quality of quick and accurate word and text reading to reading skill as a whole (Fuchs et al., 2001; Stanovich, 1980). In the current study, we represented this skill through a combination of psychometric and computerized assessments at the word and text level, and we assessed both accuracy and speed of reading. We also included nonword reading as an indicator of the ability to recognize unknown words. We included rapid object naming, because of its linkage in prior research to fluent word and text reading (Wolf & Katzir-Cohen, 2001). On all fluency-related tasks, children were found to increase their speed and accuracy from first to third grade. Further, this reading-fluency factor was an important predictor of reading comprehension. In each grade, children who had superior reading fluency had better reading comprehension. Thus, our data were consistent with automaticity views that predict that as children become quicker and more accurate, there are freed-up resources that can be used for improved comprehension, but these resources emanate mainly from superior word-reading skills and not from the special skills associated with fluent text reading per se.

Autonomy is a second aspect of automaticity accounts of the development of reading fluency. Autonomous reading has been shown to develop early in the development of reading skill. This component was indicated by interference scores obtained from the Stroop interference task. We found an interesting pattern in the relationship between this autonomous reading component and other components of our model suggesting that Stroop interference may change in its meaning between first and third grade. In first graders, although the effect was small, Stroop interference was positively related to better reading comprehension skill. In second grade, Stroop interference was related to neither reading fluency nor reading comprehension, presumably because most children were displaying reading interference. By third grade, children who were still showing large degrees of interference from the conflicting labels were likely to be less fluent readers. Further, because we identified that the interference found in the Stroop task occurred in words having high-probability GPC units, we can state that their difficulty probably stemmed from difficulties in unitizing these sublexical features.

We think that our study helps better place the autonomy component within automaticity approaches to the development of reading skill. Earlier discussions of this component viewed the uncertain relationship between Stroop interference and word fluency as problematic for automaticity views of reading (Stanovich, 1990). Autonomy was originally considered to be the by-product of a large amount of practice (Shiffrin & Schneider, 1977) and finding interference from almost the onset of reading was viewed as problematic. Recent reconceptualizations of automaticity suggest that, instead, automatic retrieval and its concomitant benefits for attentional resources accrue very quickly in learning a new skill (Logan, 1988,1997). Thus, the development of autonomous reading early in the process of learning to read is to be expected as sublexical patterns become learned and have begun to be practiced. This new-found autonomy contributes to the resource benefits that enable good reading comprehension. Thus, the current study indicates that there is a place for this component early in the acquisition of reading and that it is indicative of increasing skill.

This Stroop interference takes on a different meaning once children have achieved autonomous reading. Most children continue to develop increased fluency in recognizing new words throughout third grade (as our data indicate). By third grade, the reduced resources that word reading in general now takes allow them to use these resources in the Stroop task to inhibit the task features that distract them. However, the persistence of a difficulty in overcoming interference from irrelevant print (as indicated by large interference scores) is a sign of a basic fluency problem, which manifests as a significant negative correlation between reading fluency and autonomous reading interference in our data.

Finally, we have shown that reading fluency and automatic reading account for considerable variance in children's reading comprehension throughout the early elementary school years. However, findings also point to the diminishing role that automaticity plays in reading comprehension as children get older. The variance accounted for by predictors in the model declined steadily from first to third grade. Presumably, once most children are fluent readers, factors other than reading fluency become important for good comprehension. As noted earlier, this may be because other factors, such as general oral language skills and the ability to draw appropriate inferences, become more important as children move from learning to read to reading to learn (Chall, 1996; Nation & Snowling, 1998).

Despite the emphasis of the current study on fluency as a key skill for comprehension, we do not wish to be misunderstood as to what we believe the implications of our study are for instruction during the early grades. No matter what the age of the students, comprehension is the ultimate goal of reading and should receive emphasis instructionally throughout. At all grades, there was significant variance in children's reading comprehension *not* accounted for by fluency. Thus, if anything, our findings support the importance of carrying out fluency-oriented instruction *alongside* comprehension instruction. Further, the key is to remember that fluency is an important bridge to comprehension but not the ultimate destination. Once fluent reading has been achieved by children, fluency-oriented instruction can be phased out because, as our findings show, individual differences in comprehension begin to be increasingly related to things other than reading fluency.

In sum, the current study has found support for a simple reading fluency view in automaticity accounts of the development of fluent and automatic reading during the early elementary school years. Future research should assess the continuing developmental path of reading fluency and these other language factors in producing good comprehension as children become older.

Acknowledgments

We thank Samantha Johnson, Caroline Groff, Franklin Turner, Ann Marie Hamilton, Barbara Bradley, Matthew Quirk, and Deborah Woo for assisting in the collection of data. The project was supported by the Interagency Education Research Initiative, a program of research managed jointly by the National Science Foundation, the Institute of Education Sciences in the U.S. Department of Education, and the National Institute of Child Health and Human Development in the National Institutes of Health (NICHD/NIH). Funding for the project was provided by NICHD NIH Grant 7 R01 HD040746-06.

References

ADAMS, MJ. Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press; 1990.

ANDERSON JR. Skill acquisition: Compilation of weak-method problem solutions. Psychological Review 1987;94:192–210.

BACKMAN J, BRUCK M, HEBERT M, SEIDENBERG MS. Acquisition and use of spelling–sound correspondences in reading. Journal of Experimental Child Psychology 1984;38:114–133.

BERGEN L, GRIMES T, POTTER D. How attention partitions itself during simultaneous message presentations. Human Communication Research 2005;31:311–336.

BERNDT RS, REGGIA JA, MITCHUM CC. Empirically derived probabilities for grapheme-to-phoneme correspondences in English. Behavior Research Methods, Instrumentation, & Computers 1987;19:1–9.

BOWERS PG. Text reading and rereading: Determinants of fluency beyond word recognition. Journal of Reading Behavior 1993;25:133–153.

BOWERS PG, SWANSON LB. Naming speed deficits in reading disability: Multiple measures of a singular process. Journal of Experimental Child Psychology 1991;51:195–219. [PubMed: 2033360]

BREZNITZ Z, BERMAN L. The underlying factors of word reading rate. Educational Psychology Review 2003;15:247–265.

BYERS JA. Content and concurrent validity of the WIAT and WJ–R reading subtests for second grade students. Dissertation Abstracts International Section A: Humanities and Social Sciences 1998;58:3100.

CAIN K, OAKHILL J, BRYANT P. Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. Journal of Educational Psychology 2004;94:31–42.

CARROLL, JB.; DAVIES, P.; RICHMAN, B. The American Heritage word frequency book. New York: Houghton-Mifflin; 1971.

CASSAR M, TREIMAN R. The beginnings of orthographic knowledge: Children's knowledge of double letters in words. Journal of Educational Psychology 1997;89:631–644.

CASTEEL MA. Effects of inference necessity and reading goal on children's inferential generation. Developmental Psychology 1993;29:346–357.

CHALL, JS. Stages of reading development. Vol. 2. Fort Worth, TX: Harcourt Brace; 1996.

CLAY MM, IMLACH RH. Juncture, pitch, and stress in reading. Journal of Verbal Learning and Verbal Behavior 1971;10:133–139.

COLTHEART V, LEAHY J. Children's and adults' reading of nonwords: Effects of regularity and consistency. Journal of Experimental Psychology: Learning, Memory, and Cognition 1992;18:718–729.

COMALLI PE JR, WAPNER S, WERNER H. Interference effects of Stroop color-word test in childhood, adulthood, and aging. Journal of Genetic Psychology 1962;100:47–53. [PubMed: 13880724]

COWIE R, DOUGLAS-COWIE E, WICHMANN A. Prosodic characteristics of skilled reading: Fluency and expressiveness in 8–10 year-old readers. Language and Speech 2002;45:47–82. [PubMed: 12375819]

CRAIG HK, THOMPSON CA, WASHINGTON JA, POTTER SL. Performance of elementary grade African-American students on the Gray Oral Reading Tests. Language, Speech and Hearing Services in Schools 2004;35:141–154.

CUNNINGHAM AE, PERRY KE, STANOVICH KE. Converging evidence for the concept of orthographic processing. Reading & Writing: An Interdisciplinary Journal 2001;14:549–568.

CUNNINGHAM AE, STANOVICH KE. Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. Journal of Educational Psychology 1990;82:733–740.

CYCOWICZ YM, FRIEDMAN D, ROTHSTEIN M, SNODGRASS JG. Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. Journal of Experimental Child Psychology 1997;65:171–237. [PubMed: 9169209]

DOWHOWER SL. Effects of repeated reading on second-grade transitional readers' fluency and comprehension. Reading Research Quarterly 1987;22:389–406.

EHRI LC. Do words really interfere in naming pictures? Child Development 1976;47:502–505.

FOEGEN A, ESPIN CA, ALLINDER RM, MARKELL MA. Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. Journal of Special Education 2001;34:226–236.

FOORMAN B, FRANCIS D, FLETCHER D, LYNN A. Relation of phonological and orthographic processing to early reading: Comparing two approaches to regression-based reading-level-match designs. Journal of Educational Psychology 1996;88:639–652.

FUCHS LS, FUCHS D, HOSP MK, JENKINS JR. Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading 2001;5:239–256.

GAYAN J, OLSON RK. Genetic and environmental influences on individual differences in printed word recognition. Journal of Experimental Child Psychology 2003;84:97–123. [PubMed: 12609495]

GOLINKOFF RM, ROSINSKI RR. Decoding, semantic processing, and reading comprehension skill. Child Development 1976;47:252–258.

GOSWAMI U, ZIEGLER JC, DALTON L, SCHNEIDER W. Nonword reading across orthographies: How flexible is the choice of reading units? Applied Psycholinguistics 2003;24:235–247.

GOTTARDO A, CHIAPPE P, SIEGEL LS, STANOVICH KE. Patterns of word and nonword processing in skilled and less-skilled readers. Reading and Writing: An Interdisciplinary Journal 1999;11:465–487.

GOUGH PB. How children learn to read and why they fail. Annals of Dyslexia 1996;46:3–20.

GUTTENTAG RE, HAITH MM. Automatic processing as a function of age and reading ability. Child Development 1978;49:707–716.

HANNON B, DANEMAN M. A new tool for measuring and understanding individual differences in the component processes of reading comprehension. Journal of Educational Psychology 2001;93:103–128.

HARBER JR. Syntactic complexity: A necessary ingredient in predicting readability. Journal of Learning Disabilities 1979;12:13–19.

HASBROUCK JE, TINDAL G. Curriculum-based oral reading fluency norms for students in grades 2 through 5. Teaching Exceptional Children 1992;24:41–44.

HIEBERT EH, FISHER CW. A review of the National Reading Panel's studies on fluency: The role of text. The Elementary School Journal 2005;105:443–460.

HORN CC, MANIS FR. Development of automatic and speeded reading of printed words. Journal of Experimental Child Psychology 1987;44:92–108.

JENKINS JR, FUCHS LS, VAN DEN BROEK P, ESPIN C, DENO SL. Sources of individual differences in reading comprehension and reading fluency. Journal of Educational Psychology 2003;95:719–729.

JÖRESKOG, K.; SÖRBOM, D. LISREL 8: User's reference guide. Chicago: Scientific Software International; 1996.

JÖRESKOG, K.; SÖRBOM, D. LISREL 8.3 [Computer software]. Chicago: Scientific Software International; 1999.

JUEL C. Learning to read and write: A longitudinal study of 54 children from first to fourth grades. Journal of Educational Psychology 1988;80:437–447.

KATZIR T, SHAUL S, BREZNITZ Z, WOLF M. The universal and the unique in dyslexia: A cross-linguistic investigation of reading and reading fluency in Hebrew- and English-speaking children with reading disorders. Reading & Writing 2004;17:739–768.

KLINE, RB. Principles and practice of structural equation modeling. New York: Guilford; 2005.

KUHN MR, STAHL SA. Fluency: A review of developmental and remedial practices. Journal of Educational Psychology 2003;95:3–21.

LABERGE D, SAMUELS SJ. Toward a theory of automatic information processing in reading. Cognitive Psychology 1974;6:293–323.

LAING E, HULME C. Phonological and semantic processes influence beginning readers' ability to learn to read words. Journal of Experimental Child Psychology 1999;73:183–207. [PubMed: 10357872]

LOGAN GD. Skill and automaticity: Relations, implications, and future directions. Canadian Journal of Psychology 1985;39:367–386.

LOGAN GD. Toward an instance theory of automatization. Psychological Review 1988;95:492–527.

LOGAN GD. Automaticity and reading: Perspectives from the instance theory of automatization. Reading and Writing Quarterly: Overcoming Learning Difficulties 1997;13:123–146.

LOGAN GD, KLAPP ST. Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? Journal of Experimental Psychology: Learning, Memory, and Cognition 1991;17:179–195.

LONG DL, OPPY BJ, SEELY M. Individual differences in readers' sentence and text-level representations. Journal of Memory and Language 1997;36:129–145.

LYON GR. Toward a definition of dyslexia. Annals of Dyslexia 1995;45:3–27.

LYON GR, SHAYWITZ SE, SHAYWITZ BA. A definition of dyslexia. Annals of Dyslexia 2003;53:1–14.

MANIS FR, CUSTODIO R, SZESZULSKI PA. Development of phonological and orthographic skill: A 2-year longitudinal study of dyslexic children. Journal of Experimental Child Psychology 1993;56:64–86. [PubMed: 8366326]

MANIS FR, SEIDENBERG MS, DOI LM. See Dick RAN: Rapid naming and the longitudinal prediction of reading subskills in first and second graders. Scientific Studies of Reading 1999;3:129–157.

MARLOW A, EDWARDS RP. Test review: Gray Oral Reading Test, Third Edition (GORT–3). Journal of Psychoeducational Assessment 1998;16:90–94.

MARTENSEN H, DIJKSTRA T, MARIS E. A werd is not quite a word: On the role of sublexical phonological information in visual lexical decision. Language and Cognitive Processes 2005;20:513–552.

MCFALLS EM, SCHWANENFLUGEL PJ, STAHL S. Influence of word meaning on the acquisition of a reading vocabulary in second-grade children. Reading and Writing: An Interdisciplinary Journal 1996;8:235–250.

MORRIS RD, STUEBING KK, FLETCHER JM, SHAY-WITZ S, LYON R, SHANKWEILER D, et al. Subtypes of reading disability: Variability around a phonological core. Journal of Educational Psychology 1998;90:347–373.

NATION K, SNOWLING MJ. Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. Journal of Memory & Language 1998;39:85–101.

NATIONAL INSTITUTE OF CHILD HEALTH AND HUMAN DEVELOPMENT. Report of the subgroups: National Reading Panel. Washington, DC: U.S. Government Printing Office; 2000.

NICHOLSON, T. Reading comprehension processes. In: Thompson, GB.; Nicholson, T., editors. Learning to read: Beyond phonics and whole language. New York: Teachers College Press; 1999. p. 127-149.

NOTENBOOM A, REITSMA P. Investigating the dimensions of spelling ability. Educational & Psychological Measurement 2003;63:1039–1059.

OLSON RK, GILLIS JJ, RACK JP, DEFRIES JC, FULK-ER DW. Confirmatory factor analysis of word recognition and process measures in the Colorado Reading Project. Reading & Writing 1991;3:235–248.

OLSON, RK.; KLIEGL, R.; DAVIDSON, BJ.; FOLTZ, G. Individual and developmental differences in reading disability. In: MacKinnon, GE.; Waller, TG., editors. Reading research: Advances in theory and practice. Vol. 4. New York: Academic Press; 1985. p. 1-64.

PAAP KR, OGDEN WC. Letter encoding is an obligatory but capacity-demanding operation. Journal of Experimental Psychology: Human Perception and Performance 1981;7:518–527.

PENNINGTON BF, LEFLY DL, VAN ORDEN GC, BOOKMAN MO, SMITH SD. Is phonology bypassed in normal or dyslexic development? Annals of Dyslexia 1987;37:62–89.

PERFETTI, CA. Reading ability. London: Oxford University Press; 1985.

PERFETTI CA, HOGABOAM T. Relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology 1975;67:461–469.

PERRY C, ZIEGLER JC, COLTHEART M. How predictable is spelling? Developing and testing metrics of phoneme-grapheme contingency. The Quarterly Journal of Experimental Psychology 2002;55A: 897–915. [PubMed: 12188519]

PRITCHARD VE, NEUMANN E. Negative priming effects in children engaged in nonspatial tasks: Evidence of early development of an intact inhibitory mechanism. Developmental Psychology 2004;40:191–203. [PubMed: 14979760]

RACK JP, SNOWLING MJ, OLSON RK. The non-word reading deficit in developmental dyslexia: A review. Reading Research Quarterly 1992;27:28–53.

RASINSKI TV. Investigating measures of fluency. Educational Research Quarterly 1990;14:34–44.

RICHARDSON V, ANDERS P, TIDWELL D, LLOYD C. The relationship between teachers' beliefs and practices in reading comprehension instruction. American Educational Research Journal 1991;28:559–586.

SCARBOROUGH HS. Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. Annals of Dyslexia 1998;48:115–136.

SCHADLER M, THISSEN DM. The development of automatic word recognition and reading skill. Memory & Cognition 1981;9:132–141.

SCHNEIDER, W.; DUMAIS, ST.; SHIFFRIN, RM. Automatic and control processing and attention. In: Parasuraman, R.; Davies, R., editors. Varieties of attention. New York: Academic Press; 1984. p. 1-27.

SCHNEIDER, W.; ESCHMAN, A.; ZUCCOLOTTO, A. E-prime user's guide. Pittsburgh, PA: Psychology Software Tools; 2001.

SCHREIBER PA. On the acquisition of reading fluency. Journal of Reading Behavior 1980;12:177–186.

SCHUMACKER, RE.; LOMAX, RG. A beginner's guide to structural equation modeling. Mahwah, NJ: Erlbaum; 1996.

SCHWANENFLUGEL PJ, AKIN CE. Developmental trends in lexical decisions for abstract and concrete words. Reading Research Quarterly 1994;29:251–263.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

SCHWANENFLUGEL PJ, HAMILTON AM, KUHN MR, WISENBAKER J, STAHL SA. Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. Journal of Educational Psychology 2004;96:119–129. [PubMed: 19777077]

SCHWANENFLUGEL PJ, NOYES CR. Context availability and the development of word reading skill. Journal of Literacy Research 1996;28:35–54.

SHIFFRIN RM, SCHNEIDER W. Controlled and automatic human information processing: II. Perceptual learning, Automatic attending, and a general theory. Psychological Review 1977;84:127–190.

SHINN MR, GOOD RH, KNUTSON N, TILLY WD, COLLINS VC. Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. School Psychology Review 1992;21:459–479.

SMITH TD, SMITH BL. Relationship between the Wide Range Achievement Test 3 and the Wechsler Individual Achievement Test. Psychological Reports 1998;83:963–967. [PubMed: 9923176]

STAGE SA, SHEPPARD J, DAVIDSON MM, BROWNING MM. Prediction of first-graders' growth in oral reading fluency using kindergarten letter fluency. Journal of School Psychology 2001;39:225–237.

STANOVICH KE. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly 1980;16:32–71.

STANOVICH KE. Explaining the variance in reading Ability in terms of psychological processes: What have we learned? Annals of Dyslexia 1985;35:67–96.

STANOVICH KE. Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. Developmental Review 1990;10:72–100.

STANOVICH KE, CUNNINGHAM AE, WEST RF. A longitudinal study of the development of automatic recognition skills in first graders. Journal of Reading Behavior 1981;13:57–74.

STANOVICH KE, NATHAN RG, VALA-ROSSI M. Developmental changes in the cognitive correlates of reading ability and the developmental lag hypothesis. Reading Research Quarterly 1986;21:267–283.

STANOVICH KE, NATHAN RG, ZOLMAN JE. The developmental lag hypothesis in reading: Longitudinal and matched reading-level comparisons. Child Development 1988;59:71–86.

STANOVICH KE, SIEGEL LS. Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. Journal of Educational Psychology 1994;86:24–53.

STROOP JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology 1935;18:643–662.

TIPPER SP, BOURQUE TA, ANDERSON SH, BRE-HAUT JC. Mechanisms of attention: A developmental study. Journal of Experimental Child Psychology 1989;48:353–378. [PubMed: 2584921]

TORGESEN, JK.; WAGNER, RK.; RASHOTTE, CA. Test of Word Reading Efficiency. Austin, TX: Pro-Ed; 1999.

TRAXLER MJ. Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. Quarterly Journal of Experimental Psychology 2002;55A:75–96. [PubMed: 11873857]

TREIMAN R, GOSWAMI U, BRUCK M. Not all nonwords are alike: Implications for reading development and theory. Memory & Cognition 1990;13:357–364.

TREIMAN R, MULLENNIX J, BIJELJAC-BABIC R, RICHMOND-WELTY ED. The special role of rimes in the description, use, and acquisition of English orthography. Journal of Experimental Psychology: General 1995;124:107–136. [PubMed: 7782735]

VELLUTINO, FR.; SCANLON, DM.; TANZMAN, MS. Components of reading ability: Issues and problems in operationalizing word identification, phonological coding, and orthographic coding. In: Lyon, GR., editor. Frames of reference for the assessment of learning disabilities: New views on measurement issues. Baltimore: Paul H. Brookes; 1994. p. 279-332.

VENEZKY, RL. The American way of spelling: The structure and origins of American English orthography. New York: Guilford; 1999.

WATERS GS, SEIDENBERG MS, BRUCK M. Children's and adults' use of spelling-sound information in three reading tasks. Memory & Cognition 1984;12:293–305.

WILLOWS DM, RYAN EB. The development of grammatical Sensitivity and its relationship to early reading achievement. Reading Research Quarterly 1986;21:253–266.

WOLF M, BALLY H, MORRIS R. Automaticity, retrieval processes, and reading: A longitudinal study in average and impaired readers. Child Development 1986;57:988–1000. [PubMed: 3757613]

WOLF M, BOWERS P. The double-deficit hypothesis for the developmental dyslexias. Journal of Educational Psychology 1999;91:415–438.

WOLF M, BOWERS PG, BIDDLE K. Naming-speed processes, timing, and reading: A conceptual review. Journal of Learning Disabilities 2000;33:387–407. [PubMed: 15493099]

WOLF M, KATZIR-COHEN T. Reading fluency and its intervention. Scientific Studies of Reading 2001;5:211–239.

YOUNG A, BOWERS PG. Individual difference and text difficulty determinants of reading fluency and expressiveness. Journal of Experimental Child Psychology 1995;60:428–454.

ZENO, SM.; IVENS, SH.; MILLARD, RT.; DUVVURI, R. The educator's word frequency guide. Brewster, NY: Touchstone Applied Science Associates; 1995.

ZUTELL J, RASINSKI TV. Training teachers to attend to their students' oral reading fluency. Theory Into Practice 1991;30:211–217.

Biographies

PAULA J. SCHWANENFLUGEL is a professor of educational psychology, psychology, linguistics, and cognitive science at the University of Georgia, where she teaches courses on child development, cognition, and psycholinguistics as applied to educational settings. She has recently been engaged in grants and research on reading fluency, preliteracy skills, and vocabulary, and classroom practices related to these topics. She can be contacted at the Department of Educational Psychology and Instructional Technology, 325R Aderhold Hall, University of Georgia, Athens, GA 30602, USA, pschwan@uga.edu.

ELIZABETH B. MEISINGER is a center manager for Youth & Family Services in the Dallas Independent School District. Her research interests include the normal development of reading skills with a focus on reading fluency, the assessment of children with reading disabilities, and the integration of academic and psychological interventions for children in the public school setting. She can be contacted at Youth & Family Services, Dallas Independent School District, 2909 N. Buckner Drive, Dallas, TX 75228, USA, bmeisinger@hotmail.com.

JOSEPH M. WISENBAKER recently retired as an associate professor of educational psychology at the University of Georgia, where he taught courses in statistical methods and research design. In addition to working with colleagues across a wide range of disciplines, he has worked extensively with schools in the evaluation of educational programs. He can be contacted at Department of Educational Psychology & Instructional Technology, 325 Aderhold Hall, University of Georgia, Athens, GA 30602, USA, wisenbak@bellsouth.net.

MELANIE R. KUHN is an assistant professor in the Department of Learning and Teaching at the Rutgers Graduate School of Education, where she teaches courses on assessment and instruction for struggling readers. She has been engaged in research grants on fluency, assistive technology, and, most recently, as part of the Mid-Atlantic Collaborative for Applied Research in Education. Her research interests also include literacy instruction for struggling readers, comprehension development, and vocabulary instruction. She can be contacted at the Graduate School of Education, Rutgers University, 10 Seminary Place, New Brunswick, NJ 08901-1183, USA, melaniek@rci.rutgers.edu.

GREGORY P. STRAUSS is a doctoral student in clinical psychology at the University of Nevada–Las Vegas. He conducts research in the areas of cognition, neuropsychology, psychopathology, and reading fluency. He can be reached at the Department of Psychology, University of Nevada, Las Vegas, 4505 Maryland Parkway, Box 5030, Las Vegas, NV, USA 89154-5030, USA, straussg@unlv.nevada.edu.

ROBIN D. MORRIS is the vice president for research and Regents Professor of Psychology at Georgia State University. He also holds a joint appointment in the Department of Educational Psychology and Special Education in the College of Education. His current research is focused on reading and language development, reading disabilities and dyslexia, bilingual language and reading development, and the neuroimaging of the developing brain. He can be contacted at the Department of Psychology, Georgia State University, PO Box 5010, Atlanta, GA 30302-5010, USA, robinmorris@gsu.edu.

FIGURE 1A: TEXT READING AS MEDIATOR MODEL

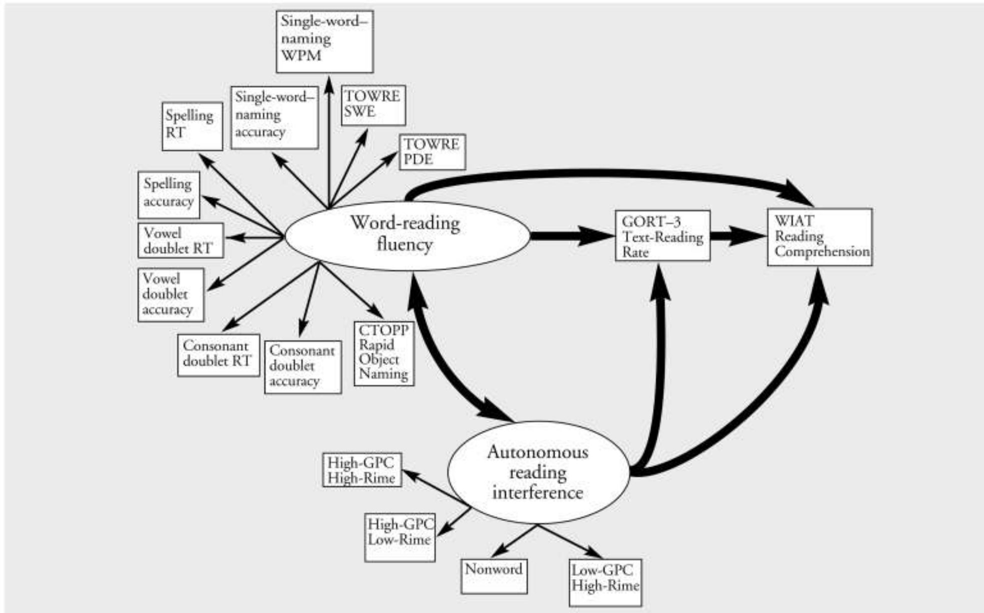

FIGURE 1B: SIMPLE READING FLUENCY MODEL

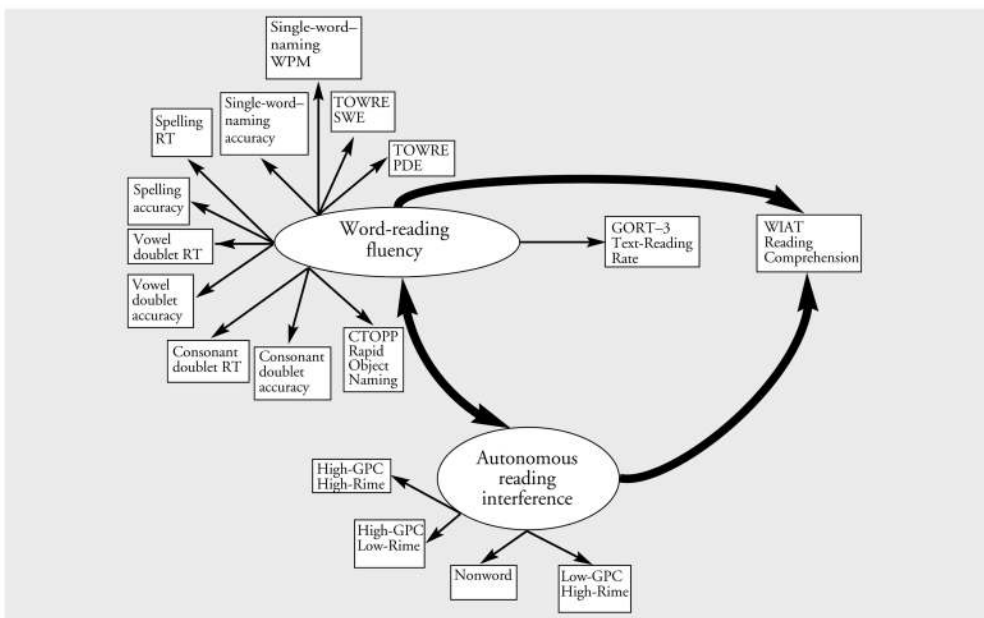

FIGURE 1.

FIGURES 1A AND B THEORETICAL MODELS FOR TWO AUTOMATICITY VIEWS OF READING FLUENCY

FIGURE 2A

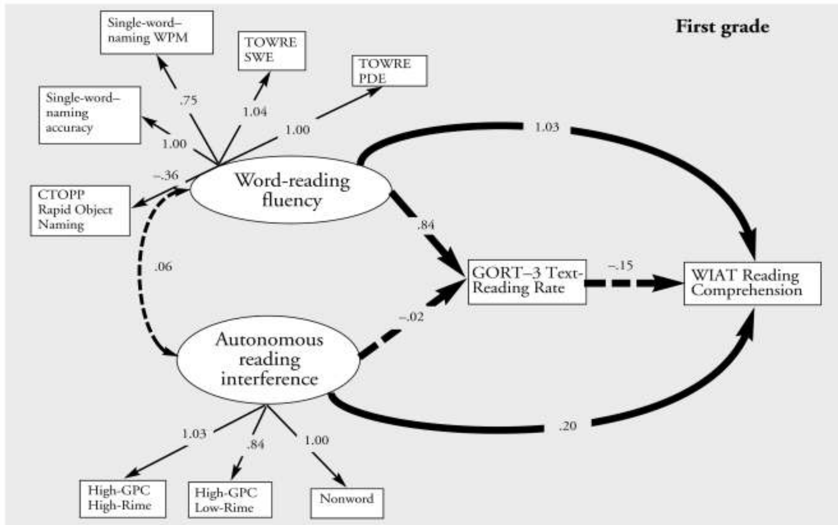

FIGURE 2B

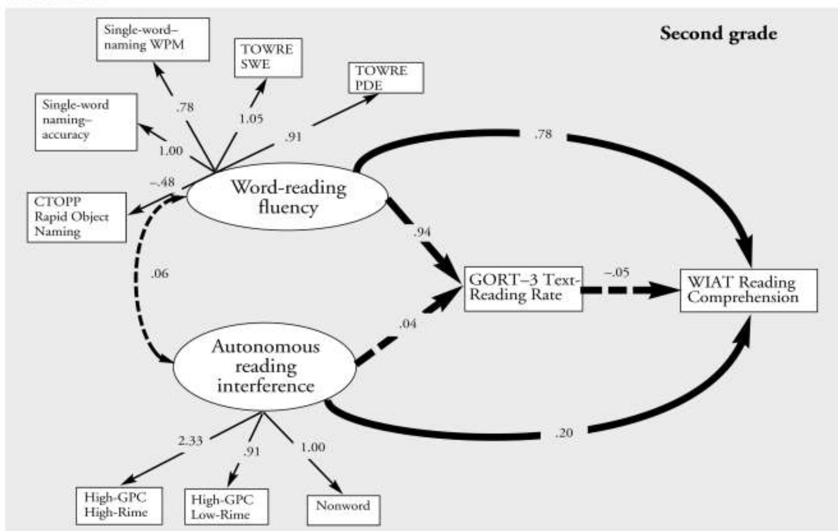

FIGURE 2.

FIGURES 2A, B, AND C TEXT READING AS MEDIATOR MODEL

Notes. Solid lines represent significant paths; dashed lines nonsignificant paths. TOWRE = Test of Word-Reading Efficiency (SWE: Sight-Word Efficiency subtest, PDE: Phonemic-Decoding Efficiency subtest); CTOPP–Rapid Object Naming = Comprehensive Test of Phonological Processing, Rapid Object Naming subtest; WIAT = Wechsler Individual Achievement Test; GORT–3 = Gray Oral Reading Test–Third Edition; Figure 2a: first grade; 2b: second grade; 2c: third grade.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

FIGURE 3A

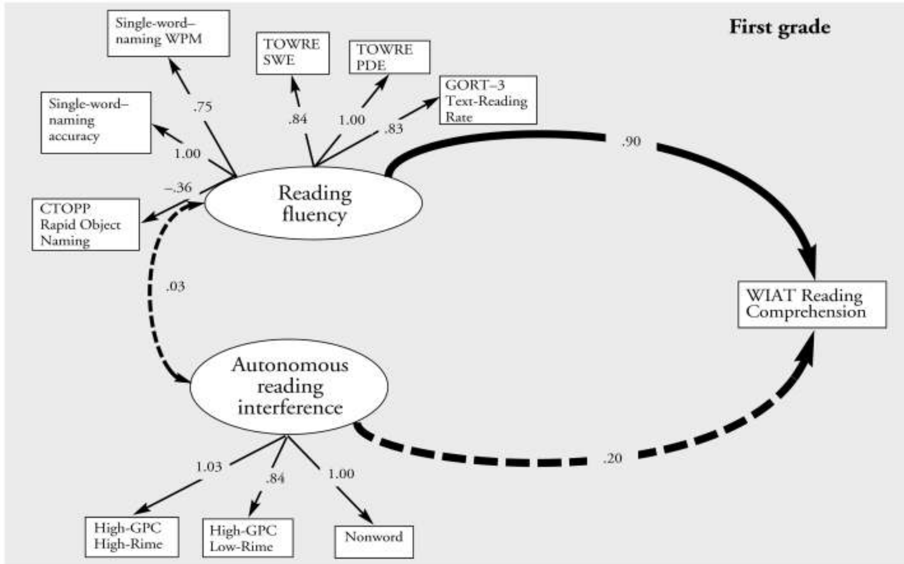

FIGURE 3B

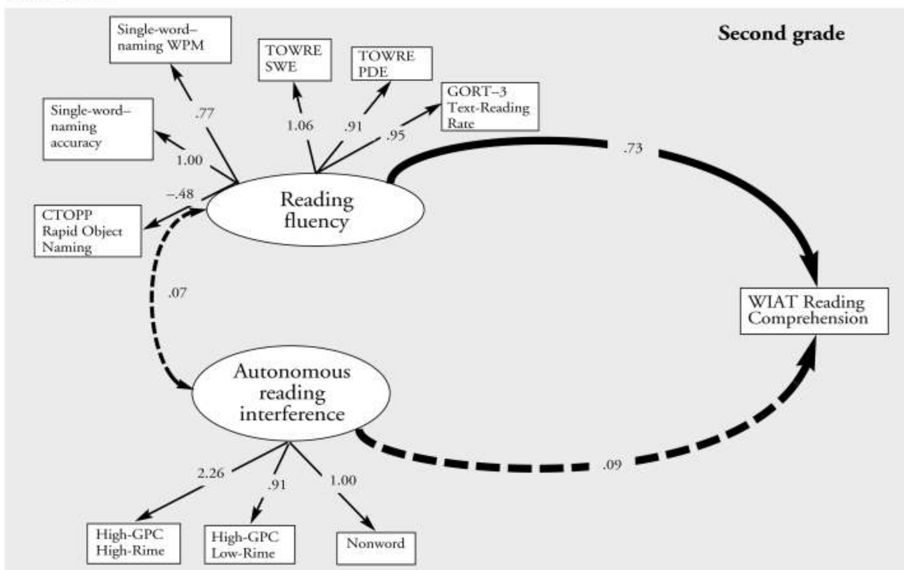

FIGURE 3.

FIGURES 3A, B, AND C SIMPLE READING FLUENCY MODEL

Notes. Solid lines represent significant paths; dashed lines nonsignificant paths. TOWRE = Test of Word-Reading Efficiency (SWE: Sight-Word Efficiency subtest, PDE: Phonemic-Decoding Efficiency subtest); CTOPP–Rapid Object Naming = Comprehensive Test of Phonological Processing, Rapid Object Naming subtest; WIAT = Wechsler Individual Achievement Test; GORT–3 = Gray Oral Reading Test–Third Edition; Figure 3a: first grade; 3b: second grade; 3c: third grade.

TABLE 1

MEAN OF THE STANDARDIZED ASSESSMENTS BY GRADE LEVEL

		Grade		
Assessment		**1**	**2**	**3**
TOWRE–SWE				
Words/45 sec.	*M*	31.88	49.05	60.44
	SD	16.17	14.37	12.04
Standard score	*M*	105	103	105
	SD	12	13	12
TOWRE–PDE				
Nonwords/45 sec.	*M*	13.15	22.25	30.68
	SD	9.89	10.95	11.17
Standard score	*M*	103	104	105
	SD	11	13	13
WIAT–RC				
No. passages correct	*M*	8.91	14.19	19.75
	SD	6.01	4.83	5.52
Standard score	*M*	96	95	103
	SD	15	11	15
CTOPP–RON				
Seconds	*M*	92	78	65
	SD	34	22	15
Standard score	*M*	9.07	8.75	9.76
	SD	3.36	2.69	2.86
GORT–3, Passage 1				
Rate sum score	*M*	2.08	5.70	13.04
	SD	3.96	5.53	7.77
Standard score	*M*	8.41	8.56	9.39
	SD	2.07	2.36	2.70

TABLE 2

MEAN RTS IN MILLISECONDS (AND ERROR RATES) FOR EXPERIMENTER-CONSTRUCTED READING TASKS

	Grade		
Task	**1**	**2**	**3**
Stroop			
Picture control	1203 (.04)	1172 (.01)	982 (.01)
Letter-string control	1192 (.04)	1274 (.03)	1029 (.03)
Nonword	1339 (.05)	1341 (.04)	1171 (.03)
Low GPC + High Rime	1231 (.07)	1266 (.05)	1148 (.04)
High GPC + High Rime	1310 (.09)	1299 (.05)	1172 (.04)
High GPC + Low Rime	1344 (.09)	1400 (.06)	1227 (.06)
Single-word naming	2231 (.73)	1314 (.42)	858 (.20)
Doublet Knowledge Task			
Vowel doublet	2261 (.29)	2421 (.12)	2122 (.07)
Consonant doublet	2164 (.33)	2143 (.26)	1878 (.28)
Experimental spelling task	2323 (.42)	2209 (.25)	1801 (.12)

TABLE 3

STANDARDIZED PATH WEIGHTS, STANDARD ERRORS, AND *T*-VALUES FOR THE TEXT READING AS MEDIATOR MODEL

Path	Weight	SE	t^{a}
First grade		Direct paths	
Word-reading fluency to			
Text-reading rate	.84	.08	10.97
Reading comprehension	1.03	.10	10.13
Autonomous reading interference to			
Text-reading rate	−.02	.10	.25
Reading comprehension	.20	.09	2.30
		Indirect path	
Text-reading rate to			
Reading comprehension	−.15	.09	1.67
Second grade		Direct paths	
Word-reading fluency to			
Text-reading rate	.94	.08	11.16
Reading comprehension	.78	.21	3.69
Autonomous reading interference to			
Text-reading rate	.04	.12	.35
Reading comprehension	.10	.16	.62
		Indirect path	
Text-reading rate to			
Reading comprehension	−.05	.18	.27
Third grade		Direct paths	
Word-reading fluency to			
Text-reading rate	1.07	.14	7.89
Reading comprehension	.72	.32	2.23
Autonomous reading interference to			
Text-reading rate	.13	.13	1.00
Reading comprehension	.18	.18	.99
		Indirect path	
Text-reading rate to			
Reading comprehension	.07	.23	.31

[a] t_{crit} (31) = 2.04, $p < .05$

TABLE 4

STANDARDIZED PATH WEIGHTS, STANDARD ERRORS, AND *T*-VALUES FOR THE SIMPLE READING FLUENCY MODEL

Path	Weight	SE	t^a
First grade		Direct paths	
Reading fluency to reading comprehension	.90	.10	10.13
Autonomous reading interference to reading comprehension	.20	.09	2.33
Second grade		Direct paths	
Reading fluency to reading comprehension	.73	.10	6.96
Autonomous reading interference to reading comprehension	.09	.16	.57
Third grade		Direct paths	
Reading fluency to reading comprehension	.79	.15	5.27
Autonomous reading interference to reading comprehension	.17	.17	1.00

[a] t_{crit} (31) = 2.04, $p < .05$