American Educational Research Journal
March 2007, Vol. 44, No. 1, pp. 161–196
DOI: 10.3102/0002831206298171

The Impact of Science Knowledge, Reading Skill, and Reading Strategy Knowledge on More Traditional "High-Stakes" Measures of High School Students' Science Achievement

Tenaha O'Reilly
Educational Testing Service
Danielle S. McNamara
University of Memphis

*This study examined how well cognitive abilities predict high school students' science achievement as measured by traditional content-based tests. Students (*n *= 1,651) from four high schools in three states were assessed on their science knowledge, reading skill, and reading strategy knowledge. The dependent variable, content-based science achievement, was measured in terms of students' comprehension of a science passage, science course grade, and state science test scores. The cognitive variables reliably predicted all three measures of science achievement, and there were also significant gender differences. Reading skill helped the learner compensate for deficits in science knowledge for most measures of achievement and had a larger effect on achievement scores for higher knowledge than lower knowledge students. Implications for pedagogy and science assessment are discussed.*

Keywords: science achievement, prediction, cognitive ability, gender, compensation

In recent years there has been a growing concern regarding the current state of the educational system in the United States. For example, in comparison to other countries, high school students in the United States are falling behind students in other countries on various measures of academic achievement and, in particular, on measures of reading comprehension (Perie, Grigg, & Donahue, 2005; Snow, 2002). To compound the problem, teachers rarely provide instruction on strategies that emphasize comprehension (e.g., Pressley, 2002; Taylor, Pearson, Clark, & Walpole, 1999), and students who have comprehension difficulties rarely engage in such strategic processing (Lenski & Nierstheimer, 2002).

Given these concerns, one goal of this research was to examine the relative contributions of cognitive abilities to students' science achievement. By *science achievement*, we are referring to more traditional content-based assessments of science achievement such as those typically found on high-stakes tests, in contrast to more "authentic" (see Chinn & Malhotra, 2002) or practitioner-based approaches to assessment. We chose to examine content-based science achievement because content-based assessments tend to dominate many high-stakes tests such as state standards tests. In turn, whereas these high-stakes tests may utilize "nonoptimal" ways to assess science achievement (Chester, 2005a, 2005b), they continue to affect education through their impact on important policy decisions (Kane, 2002). In this study, we examined how cognitive abilities affect scores on such content-based assessments of science achievement.

Prior research has shown the importance of domain knowledge (e.g., Dochy, Segers, & Buehl, 1999), reading skill (e.g., Voss & Silfies, 1996), and reading strategy knowledge (e.g., Cottrell & McNamara, 2002; O'Reilly & McNamara, 2002a) for science comprehension. The overarching goal of this work was not only to discern how these student abilities affect high school students' content-based science achievement but also to determine whether these cognitive abilities can partially compensate for one another. Although the individual effects of these factors on learning have been examined in separate studies, to the best of our knowledge no study has simultaneously measured the impact of all three variables on students' comprehension and achievement with a wide range of students.

The focus of this study is on students in high school science classes. For a diverse sample, the study includes students from suburban, rural, and urban schools including a variety of socioeconomic and ethnographic backgrounds. Our goal was to examine which cognitive abilities most contribute to a science student's achievement in a science course. Thus, we compared the cognitive measures' relationship to three types of outcome measures: (a) students' ability to comprehend a typical science text; (b) their achievement in a science course as measured by course grades; and (c) for two schools, the students' achievement on a statewide measure of students' science achievement (i.e., Virginia's Standards of Learning or state science test).

Tenaha O'Reilly is an associate research scientist in the Center of Validity Research, Educational Testing Service, Rosedale Road, MS 10-R, Princeton, NJ 08541; e-mail: *toreilly@ets.org*. This work was conducted while he was a postdoctorate fellow at the University of Memphis. He is interested in investigating ways to improve reading comprehension, critical thinking, and the transfer of training through the use of technology and models of human expertise.

Danielle S. McNamara is an associate professor in the Department of Psychology, University of Memphis, Psychology Building, Rm. 434, Memphis, TN 38152; e-mail: *d.mcnamara@mail.psyc.memphis.edu*. She investigates reading comprehension, memory, and learning and develops technologies to improve learning from text.

The Role of Knowledge

Our focus on science knowledge, reading skill, and reading strategies was motivated by a long history of research showing the importance of these abilities to comprehension and learning. Research has shown that a reader's domain knowledge can have a dramatic impact on how well new information is comprehended and learned (Alexander, 1997; Bransford, Brown, & Cocking, 1999; Dochy et al., 1999; Shapiro, 2004; Thompson & Zamboanga, 2004; Willoughby, Waller, Wood, & MacKinnon, 1993; Willoughby, Wood, & Khan, 1994). In short, knowledge is the foundation with which new information is learned. For example, information that is more easily integrated with knowledge is often remembered better than material that is more difficult to integrate with existing knowledge (Pressley et al., 1992; Willoughby et al., 1993; Woloshyn, Schnieder, & Pressley, 1992). Similarly, readers with greater domain knowledge are more likely to remember the main ideas of a text (Spilich, Vesonder, Chiesi, & Voss, 1979), use effective reading strategies (Lundeberg, 1987), and convey greater interest in the reading material than lower knowledge readers (Tobias, 1994; Zhang & Zhang, 1996). Other work has shown that the effectiveness of various forms of instruction (Mayer, 2001) and certain learning strategies (Willoughby et al., 1994) often depend upon the level of the learner's domain knowledge. For example, higher knowledge learners are better able to compensate for poor instructional formats (Mayer, 2001) and are more proficient in using learning strategies that require elaboration (Willoughby et al., 1994). In sum, a learner's existing knowledge has a large impact on knowledge acquisition.

In terms of academic achievement, the role of domain knowledge is probably most critical for helping students to interpret and comprehend their textbooks. For instance, some researchers have argued that learners often make more errors on various tasks as a result of *missing* knowledge, rather than *incorrect* knowledge (VanLehn, 1990, 1995). More specifically, many school texts are difficult to understand because they often omit important background information and fail to make relations among concepts in the text explicit (Beck, McKeown, & Gromoll, 1989; Chi, De Leeuw, Chiu, & LaVancher, 1994; VanLehn, 1996, 1998; Wilson & Anderson, 1986). Accordingly, domain knowledge can facilitate textbook comprehension by providing the reader with the necessary resources to fill in conceptual gaps (McNamara, 2001; McNamara & Kintsch, 1996; McNamara, Kintsch, Songer, & Kintsch, 1996). Increased knowledge provides a framework for readers to draw the critical inferences required to make sense of incomplete textbooks. On a more general plane, effective learning is much more than simply recalling or recognizing pieces of a text; comprehension involves the construction of a highly structured representation of the text in *relation* to the learner's knowledge and experience (McNamara & Kintsch, 1996; Snow, 2002). In other words, learning is optimal when the reader actively integrates the information contained in the text with the reader's relevant background knowledge.

Collectively, these studies suggest that the learner's background knowledge is important in determining how well readers can comprehend, learn, and utilize new information (Bransford et al., 1999).

Although a student's knowledge plays a vital role in learning, comprehension, and academic achievement, having the knowledge available does not guarantee that the student will use the knowledge, or use the knowledge appropriately (Ross, 1989; Ross & Bradshaw, 1994; Ross & Kennedy, 1990). Students often fail to rely on their knowledge to help them comprehend new material (Bransford et al., 1982; Wood, Pressley, & Winne, 1990), despite the potential benefits of utilizing such knowledge. To compound the issue, students who have comprehension difficulty often fail to employ strategies that are necessary to improve comprehension at a deeper level (Lenski & Nierstheimer, 2002). Failure to employ effective strategies causes the student to process the information "mindlessly" (Langer, 1989) and settle for only a superficial understanding of the material (Graesser, Person, & Harter, 2001; Pressley et al., 1992).

The Impact of Reading Skill and Metacognitive Reading Strategies

One important cognitive ability that may have an impact on the effective use of knowledge is reading skill (see Snow, Burns, & Griffin, 1998). Reading skill is broadly defined as the ability to develop a coherent representation of the text that matches the intended message to the reader. Of course, research has demonstrated the importance of reading skill for both comprehension (e.g., Voss & Silfies, 1996) and academic achievement (e.g., Alcock et al., 2000). Although there are many factors that contribute to reading skill, there are two attributes of skilled readers that are of particular interest in the current study: inference making and reading strategy knowledge. For example, one of the hallmarks of skilled reading is the strategic ability to use knowledge to generate inferences (e.g., Long, Oppy, & Seely, 1994; Magliano & Millis, 2003; Magliano, Wiemer-Hastings, Millis, Muñoz, & McNamara, 2002; Oakhill, 1984; Oakhill & Yuill, 1996). In general, skilled readers are better able to generate inferences than less skilled readers in many situations, such as solving anaphoric references, selecting the meaning of homographs, processing garden-path sentences, making appropriate inferences on line, and integrating text structures (e.g., Long & Golding, 1993; Long et al., 1994; Oakhill, 1983, 1984; Singer, Andrusiak, Reisdorf, & Black, 1992; Singer & Ritchot, 1996; Whitney, Ritchie, & Clark, 1991; Yuill & Oakhill, 1988). In other words, skilled readers are better able to use their knowledge to fill in the conceptual gaps of the text (e.g., McNamara & Kintsch, 1996; Oakhill, 1984; O'Reilly & McNamara, 2002b; Yuill & Oakhill, 1988).

Research indicates that skilled readers generate more inferences than less skilled readers because skilled readers are more strategic and have more knowledge of reading strategies. They are more likely to monitor their comprehension and use active reading strategies such as previewing, predicting, making inferences, drawing from background knowledge, and summarizing. Many researchers have argued that the key difference between skilled and less skilled readers resides in the ability to use reading strategies and knowledge in effective ways (e.g., Bereiter & Bird, 1985; Ericsson & Kintsch, 1995; MacDonald & Christiansen, 2002; McNamara, 1997; McNamara, de Vega, & O'Reilly, in press; McNamara & McDaniel, 2004; McNamara & Scott, 2001; Paris, Cross, & Lipson, 1984; Paris & Jacobs, 1984; Pressley & Afflerbach, 1995; Snow et al., 1998). This claim is based on findings that skilled readers use reading strategies (Bereiter & Bird, 1985; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Paris & Jacobs, 1984; Pressley & Afflerbach, 1995), which improves comprehension (Chi et al. 1989; Long et al., 1994; Oakhill, 1984; Oakhill & Yuill, 1996), and reading strategy training improves comprehension (Chi et al., 1994; McNamara, 2004; Palincsar & Brown, 1984; Paris et al., 1984; Pressley et al., 1992). We acknowledge that there are limits to strategy training such as possible low transfer to new domains (Brown, Bransford, Ferrara, & Campione, 1983; Garner, 1990; Saloman & Perkins, 1989) and also that some strategies are more appropriate for certain domains than others (Alexander & Judy, 1988). On balance, however, the benefits of reading comprehension strategies have been well established in a variety of studies (Chi et al., 1994; McNamara, 2004; Pressley et al., 1992; Rosenshine & Meister, 1994) and, consequently, we predict that they are useful across many areas of science.

The Role of Inquiry Learning and Authentic Approaches to Science Education

Current trends in science education have called for a significant departure from traditional content-based approaches of science instruction to more of a practitioner model. For example, the National Research Council (1996) has set inquiry learning as a curriculum goal for students in Grades 2 to 12. These standards place emphasis on learning science by engaging in the type of scientific enterprise that real scientists conduct, such as formulating questions, designing experiments, controlling for variables, taking measurements, and drawing conclusions based upon evidence. The goal is to create a curriculum that focuses on deep conceptual understanding of science, as opposed to a collection of superficial fact-based concepts (Kesidou & Roseman, 2002). Furthermore, researchers have argued that many of the current science texts are inadequate because they contain misconceptions and errors (e.g., Roseman, Kesidou, Stern, & Caldwell, 1999), teachers claim that they are often outdated (Timmerman, 2002; National Education Association [NEA], 2002), and they often fail to capture the process of real or authentic science (Chinn & Malhotra, 2002). Chinn and Malhotra (2002) argued that the process of inquiry described in traditional science texts is not an accurate representation of what real scientists do. For example, traditional science texts often depict science inquiry as a linear sequence of steps. These depictions largely ignore that seeking science explanations is recursive and that scientists often use heuristics as they build and revise theoretical models.

Given these inherent problems with current science texts, it is not surprising that many teachers are reluctant to use science texts in their classrooms (Driscoll, Moallem, Dick, & Kirby, 1994), and in some project-based schools, texts are not used at all (e.g., Radcliffe, Caverly, Peterson, & Emmons, 2004). Despite these problems, however, studies indicate that many science teachers do use textbooks in their instructional practices (Lumpe & Beck, 1996). For instance, the NEA 2002 survey indicated that 47% of teachers reported using a textbook every day as a part of their classroom instruction. Nonetheless, whereas many teachers report using textbooks, it is not entirely clear how they are using the text for instruction. For instance, Driscoll et al. (1994) found that teachers often use texts as a dictionary or a resource that students can consult to look up difficult scientific terms. The idea of teaching using texts as a reference was also supported by the NEA report: Ninety-five percent of teachers indicated that they used texts as a reference tool, 91% indicated they used texts for supplement planning, 64% for homework, 89% as a guide for student lessons, and 80% for class debate and discussion (NEA, 2002). Thus, although there is certainly a wide array of ways in which textbooks are used in the classroom, they continue to play an important role in teacher instruction (Weiss, Pasley, Smith, Banilower, & Heck, 2003).

In sum, teaching practices in science are now changing from the more fact-based or content delivery method to a practitioner or "science as inquiry" model. Second, there are a number of problems with current textbooks, and there is a large variability in the manner in which teachers use textbooks. These issues might seem to suggest that content learning and assessment now play a minor role in education. Bransford et al. (1999) have shown in their seminal work, *How People Learn*, that learner-centered inquiry should include the building of a classroom learning community that values both content knowledge and authentic forms of assessment. However, it is still important to examine the impact of traditional content-based approaches to science achievement because high-stakes tests continue to rely on more narrow and fact-based assessments of science achievement (Chester, 2005a, 2005b; Lane, 2004), and in turn, the outcomes of these tests influence policy (e.g., Carnoy, 2005; Hess, 2005; Kane, 2002). Consequently, we chose to assess how the cognitive abilities affect achievement as measured by more traditional high-stakes measures of science achievement.

The Current Study

Research has shown that reading skill and effective reading strategies help readers to use their knowledge to draw inferences and successfully understand text. This study further investigates this notion by looking at multiple measures of high school students' science comprehension and achievement and their relation to knowledge, reading skill, and reading strategies. To assess science knowledge, we gave the students a multiple-choice test consisting of questions on scientific methods, mathematics, earth science, and meteorology. A subset of this test was taken from a measure of science knowledge used in several studies (O'Reilly, Best, & McNamara, 2004; O'Reilly & McNamara, 2002b; Ozuru, Best, & McNamara, 2004). The test was designed as a baseline measure of knowledge that was relevant to the domain under our investigation, earth science. Questions on earth science and meteorology were included because the topic of the target text used to assess science comprehension was from the same domain (i.e., air masses and weather patterns). The type of questions included in this measure of science knowledge is more like the content-based questions typically found on high-stakes tests. In fact, some of the questions were taken directly from state standards tests. Thus, these questions differ from more authentic measures of science achievement that require inquiry skills and scientific reasoning.

We used a traditional measure of reading skill, the Gates-MacGinitie reading skill test (MacGinitie & MacGinitie, 1989) to gauge students' overall reading ability. A measure called the Metacomprehension Strategy Index (Schmitt, 1990) or "strategy knowledge" test was used to assess students' knowledge of metacognitive reading strategies. By *metacognitive reading strategies*, we mean students' awareness of metacognitive reading strategies as it specifically pertains to reading comprehension, as opposed to a general or all-encompassing form of self-regulated learning. The strategy knowledge test is a 25-item multiple-choice questionnaire that assesses students' knowledge of effective strategic reading processes to be employed before, during, and after reading a text.

Can Reading Skill Compensate for Low Knowledge?

One goal of this study is to examine whether these cognitive abilities can partially compensate for one another (e.g., Adams, Bell, & Perfetti, 1995; Perfetti, 1989; Walczyk, Marsiglia, Bryan, & Naquin, 2001; Walker, 1987) or if one ability in particular is necessary for students to comprehend and learn scientific information. Specifically, we examined whether reading skill or metacognitive reading strategies help students to partially compensate for deficiencies in terms of science knowledge. By "partial compensation," we mean that a decrement in performance on an achievement measure (e.g., science passage comprehension) caused by a relative deficiency in one ability (e.g., low knowledge) can be reduced or eliminated by strengths in other abilities (e.g., reading skill). Given the wealth of research on the positive effects of domain knowledge on comprehension (Shapiro, 2004), under most circumstances, it is expected that a lower knowledge student would score well below a higher knowledge student on measures of achievement that are related to domain knowledge. Let us assume that on a typical measure of passage comprehension, higher knowledge readers score on average 20% higher than lower knowledge readers. By partial compensation, we mean that a typical decrement (in the example 20%) can be greatly reduced (e.g., 10% difference between high and low knowledge) or eliminated (0% difference) if the student is proficient in another area (e.g., reading skill or metacognitive knowledge).

Similar with this logic, some studies have indicated that domain knowledge can compensate for low reading ability (e.g., Chiesi, Spilich, & Voss, 1979; Walker, 1987). Students are often not equipped, however, with the information or knowledge necessary to succeed in their courses (Snow, 2002). In turn, authors tend to leave out information when writing texts because they assume (sometimes inappropriately) that readers already have a certain level of background knowledge (Beck et al., 1989; Chi et al., 1994; VanLehn, 1996, 1998). One question that we address here is whether reading skill or knowledge of reading strategies can partially compensate for knowledge deficits in science. Given that skilled readers are more likely to generate inferences, and more metacognitive readers are more likely to monitor their comprehension and use reading strategies, then it is probable that both reading skill and metacognitive knowledge can help offset the knowledge deficits. The increased inferencing ability and strategy use may enable readers to make the inferences necessary to bridge some gaps within a text. In turn, the ability to make sense of texts with cohesion gaps may allow lower knowledge readers to comprehend the text as well, or nearly as well, as higher knowledge readers (e.g., McNamara, 2004).

Additional Benefits of Reading Skill for Higher Knowledge Readers?

We also examine whether reading skill or metacognitive reading strategy knowledge has effects on science achievement that go beyond the benefits of domain knowledge. One possibility is that knowledge may account for the majority of the variance for the science achievement measures. For example, Shapiro (2004) found with college students that reading skill accounted for very little of the variance in predicting passage comprehension as compared to knowledge. Reading skill may result in greater memory resources and more efficient search of knowledge (e.g., Ericsson & Kintsch, 1995). In addition, reading skill may enable higher knowledge readers to generate more knowledge-based inferences. Comprehension would be enhanced if knowledge of strategies allows appropriate information to be retrieved and utilized in appropriate situations (e.g., when there are conceptual gaps in the text). If this is the case, we should find additional benefits of reading skill or reading strategy knowledge over and above the effects of domain knowledge.

Gender Differences

As a final aim, we were also interested in examining whether there were any gender differences among our measures of science achievement and whether the relationship between our measures of cognitive abilities and science achievement were mediated by gender. Although gender differences in academic achievement have been shrinking or even reversed in recent years (see Fergusson & Horwood, 1997; Myers, 2002), many studies continue to find gender differences. For example, studies have shown that males are more likely to score higher than females on measures of science achievement (Kahle, 2004; Mau & Cheng, 2000), and males are more likely to be represented in samples of the highest achieving students in mathematics and science (Reis & Park, 2001). Conversely, research has shown that females score higher than males in reading comprehension (Mau & Cheng, 2000), and males are more likely to have reading disabilities (Flannery, Liederman, Daly, & Schultz, 2000; Liederman, Kantrowitz, & Flannery, 2005). Other work has indicated that females are more likely to use overt study strategies than males (Slotte, Lonka, & Lindblom, 2001).

As a result of these and other findings, we predicted that females would score higher than males on measures of reading skill and knowledge of reading strategies, whereas males were expected to perform higher than females on measures of science knowledge (e.g., prior science knowledge, state science test). It was not clear as to whether males or females would score higher on the passage comprehension. Females may be expected to score higher because they are more skilled readers, and males may be expected to score higher because males are more likely to score higher on science content.

Method

Participants

The sample consisted of 1,651 high school students from four schools: 498 students were from a high school in Norfolk, Virginia; 372 were from a high school in Americus, Georgia; 364 were from a high school in Prestonsburg, Kentucky; and 417 were from a high school in Williamsburg, Virginia. Students' grade level ranged from 9 to 12, and the average age of the students was 16.25 years (min = 13.5; max = 20.33). The students were enrolled in one of five possible class domains including biology (n = 687), chemistry (n = 222), earth science (n = 380), physical science/biology (n = 122), and physical science (n = 240).

Table 1 lists the various details of the schools including percentage of free lunches. As can be seen from the table, the sample includes a wide variety of school locales, ethnic makeup, and socioeconomic status (SES). There were three main localities included in the study: inner city, rural, and suburban, and the ethnic composition varied from primarily White (98% in rural Kentucky school; 73% in suburban Virginia school) to primarily African American (64% inner-city Virginia; 60% rural Georgia). In terms of SES of the schools, we included two measures. First, we included self-report data that were collected from the students; this median family income measure suggests that the suburban Virginia school had an SES higher than the other three schools. Second, as a more objective measure of SES, we also obtained the percentage of students in each school who were on free or reduced

Table 1
Characteristics of the Four Schools Included in the Study

	N	Location	Locale	% White	% School Free/ Reduced Lunch (*n*)	Median Family Income	Mean Age
School 1	498	Norfolk, VA	Inner city	24	51 (765)	$41-60,000	16.1
School 2	372	Sumter, GA	Rural	28	100 (620)	$41-60,000	16.4
School 3	364	Prestonsburg, KY	Rural	98	39 (257)	$41-60,000	16.3
School 4	417	Williamsburg, VA	Suburban	73	4 (56)	$61-80,000	16.2

lunch programs. The highest percentage of free or reduced lunches was found in the rural Georgia school (100%), and the lowest percentage was found in the suburban Virginia school (4%). Although the sample did not include all types of schools and regions in America, the heterogeneous nature of the sample encompassed a wide range of schools.

Design

Science achievement was treated as the dependent measure and was assessed in terms of comprehension of a science passage, students' science course grades, and a statewide measure of students' science achievement (Virginia's Standards of Learning). Our independent variables included science knowledge, reading skill, and knowledge of metacognitive reading strategies.

Materials

Science knowledge test. Science knowledge was measured with a four-response, multiple-choice test on science information (see the appendix). The test was intended to reflect the types of questions that are typically found on content-based high-stakes tests and not deeper authentic measures of science knowledge. The test consisted of questions concerning scientific methods, mathematics, earth science, and meteorology. Questions on scientific methods and mathematics were selected because both areas are common to most areas of sciences. Questions on earth science and meteorology were included because they were from the same domain as our science comprehension passage (i.e., air masses passage described below). Although some of the knowledge questions were from the same domain as the passage, none of the knowledge questions could be answered from the passage and vice versa.

The current 18-item measure was derived from a larger 48-item science knowledge measure developed in our laboratory. The 48-item measure included questions on scientific tools (e.g., optical, microscope, protective clothing), forms of energy (e.g., visible light, solar energy, geothermal power), space (universe, meteorites, constellations), scientific inquiry (test, hypothesis, explore, classify, observe), earth features (caverns, biosphere, streams, surface, resource), and mathematics (range, value, decimals, formula, percentage). Questions from the 48-item measure were drawn from a variety of sources including state tests, science textbooks, and a pool of experimenter-generated items. Specifically, questions were drawn from the following sources: Addison-Wesley, GED Guide, Georgia Guide, Holt Life Science Kentucky Standards of Learning, Merrill Physical Science, Nation's Report Card, Science Bowl, and the 2000 edition of the Virginia Standards of Learning Assessment tests for science. Pilot tests revealed that the 48-item measure robustly predicted science passage comprehension. The test was reduced to 18 items for this study due to time constraints (i.e., the tests were delivered during students' class periods). Cronbach's alpha for the 18-item science knowledge measure was $\alpha = .74$ using the Spearman-Brown formula adjusted for 30 items (see Lord & Novick, 1968).

Reading skill test. Reading skill was measured by the Gates-MacGinitie (3rd ed.) reading skill test (form L) level 7/9 (MacGinitie & MacGinitie, 1989). The test consisted of 48 multiple-choice questions designed to assess student comprehension on several short text passages. We were solely interested in reading comprehension and not vocabulary scores; therefore, the vocabulary section of the test was not administered. Reliability for the Gates-MacGinitie reading skill test in this study was $\alpha = .95$.

Metacognitive reading strategy knowledge. Metacognitive reading strategy knowledge was measured by the Metacomprehension Strategy Index (Schmitt, 1990). The Metacomprehension Strategy Index, or "strategy knowledge" test, is a 25-item multiple-choice questionnaire designed to measure knowledge of metacognitive reading strategies such as predicting and verifying, previewing, purpose setting, self-questioning, drawing from background knowledge, and summarizing. The original scale was developed to be used with narrative passages; however, the scale was adapted for use with expository passages (see Forget, 1999). In this study, the Cronbach's alpha for the strategy knowledge test was $\alpha = .72$ using the Spearman-Brown formula adjusted for 30 items. An example strategy knowledge question is "Before I begin reading, it's a good idea to a) see how many pages are in the reading; b) look up all of the big words in the dictionary; c) make some guesses about what I think the reading is about; and d) be sure I can answer the questions at the end of the last chapter."

Science comprehension passage. Finally, participants were given an 840-word passage on meteorology that covered the types and origins of air masses as well as their impact on weather patterns. The passage was taken from chapter 29, titled "Air Masses and Fronts," in Heath's Earth Science text (Spaulding & Namowitz, 1994). Heath's Earth Science is often studied in the context of a Grade 10 or 11 course on earth science. The passage included only the first four topics of the chapter. It contained 74 sentences and had a Flesch Reading Ease of 70.63 and a Flesch-Kincaid Grade Level of 6.22. The earth science domain was chosen because it is a representative science course many students take in high school, and it is not as heavily dependent on mathematics skills as other sciences (e.g., physics, chemistry). Furthermore, we chose the topic of meteorology because weather is a common concept that affects everyone; thus, it is not as abstract and unfamiliar as other science concepts (e.g., electrons, plasma, black holes).

The researchers created 20 questions based on the information contained in the air mass chapter. A variety of question response formats and question categories were created to ensure the questions could assess comprehension on a variety of levels. Three types of questions were created: text-based, bridging inference, and mental model. Text-based questions assessed local processing and thus required little or no inferencing; the answers to these questions were found in a single sentence. Bridging inference questions required the reader to draw inferences, and the answers to these questions were contained in more than one sentence. Mental model questions concerned a major theme or idea in the passage and required the reader to have a global understanding of the passage; consequently, the answers to mental model questions were found in multiple sentences. In addition, two types of question formats were created, open-ended and multiple-choice. Although both question formats assess comprehension, the open-ended format is more closely associated with recall, whereas the multiple-choice questions are likely to tap into recognition and reasoning processes.

Based on the above question classification scheme, 8 multiple-choice (4 text-based, 4 bridging) and 12 open-ended (4 text-based, 4 bridging, 4 mental model) comprehension questions were created for the passage. Distracters for the multiple-choice questions were systematically chosen based on their graded association with the target; this process is similar to the Graesser and Wisher (2001) guidelines for generating multiple-choice questions. Cronbach's alpha for the open-ended questions was $\alpha = .87$, whereas alpha level for the multiple-choice questions was $\alpha = .83$ using the Spearman-Brown formula adjusted for 30 items. An example of a multiple-choice question is "If an air mass is humid, or the surface is warm water, then a) the air mass will pass through quickly; b) the air mass will become cooler; c) showers may form; d) the weather stays fair." An example of an open-ended question is "What are the characteristics of a continental polar (cP) air mass?"

Design and Procedure

The students were tested during regular classroom hours in a 90-minute class period, or two 50-minute class periods, and all testing was conducted near the end of the academic year. The complete set of materials was presented in a single booklet with "stop" pages inserted between each measure. If students finished a particular test early, they could recheck their answers but could not go on to the next section. The participants completed the measures in the following order and time frame: science passage and questions

(20 minutes), reading skill test (20 minutes), science knowledge test (10 minutes), and strategy knowledge test (10 minutes). The science knowledge test was given after the science passage so as to not prime any concepts related to the science passage. Priming is of importance in this experiment because it may mitigate some of the effects of reading skill and metacognitive strategies. If we assume that comprehension skill plays a role in regulating the access and use of relevant knowledge and reading strategies (see McNamara, 1997; McNamara et al., 1996; McNamara & Kintsch, 1996; McNamara & O'Reilly, in press), then presenting additional cues to participants (administering related knowledge test) may weaken or reduce the impact of reading skill. In any event, none of the questions in either the science knowledge test or the passage comprehension could be potentially answered from the other.

At the end of the academic year, the students' science course grade and their Standards of Learning science scores were collected. The students' raw course grades were standardized using a *z* score transformation. The Standards of Learning test was a Virginia state test; therefore, only the students in the Norfolk and Williamsburg schools received scores on this measure of achievement.

Results

Overall Means and Effects of Gender

The following results were significant at the $p < .01$ level unless noted otherwise. It was verified for all analyses reported here that students' age differences did not alter the pattern of results. Table 2 lists the means and standard deviations of our independent and dependent measures. As can be seen from the table, our measures contain sufficient variation for analysis. Table 3 lists the means, standard deviations, and effect sizes (Cohen's *d*) for our independent and dependent measures as a function of gender. For all measures, there were significant differences as a function of gender at the $p < .01$ level, with the exception of the effect of gender on reading skill, which was significant at the $p = .017$ level. Overall, males scored significantly higher than females on measures of science knowledge, state science test, and the passage comprehension measures. Females, on the other hand, scored significantly higher than males on measures of strategy knowledge, reading skill, and final course grade. However, it should be noted that the gender differences were relatively small overall; the average effect size in terms of Cohen's *d* was .19 ($SD = .08$) across all our measures. The largest difference was observed on the Standards of Learning (Cohen's $d = .34$), whereas the smallest difference appeared on the open-ended comprehension questions (Cohen's $d = .08$).

Effects of Grade Level

Table 4 lists the measures as a function of grade level. Unfortunately, 25% of the students failed to report their current grade level, so we performed

Table 2
Means and Standard Deviations of the Proportion Correct for Independent and Dependent Variables Used in This Study

Measure	*N*	Min	Max	*M*	*SD*
Reading strategy knowledge	1,561.00	0.04	1.00	0.43	0.16
Science knowledge	1,487.00	0.00	0.94	0.44	0.17
Reading skill (number correct)	1,504.00	0.00	47.00	21.47	10.18
State science test	744.00	312.00	597.00	423.22	43.05
Course grade (standardized)	1,598.00	–3.41	2.52	0.00	0.97
Multiple-choice comprehension	1,492.00	0.00	1.00	0.41	0.24
Open-ended comprehension	1,383.00	0.00	0.63	0.13	0.12

Table 3
Means and Standard Deviations of the Proportion Correct (for Reading Skill Number Correct) for Independent and Dependent Variables Used in This Study as a Function of Gender

Measure	Gender	*n*	Min	Max	*M*	*SD*	Cohen's *d*
Strategy knowledge	Male	761.00	0.04	0.96	0.42	0.15	0.19
	Female	800.00	0.08	1.00	0.45	0.16	
Science knowledge	Male	729.00	0.00	0.89	0.46	0.18	0.23
	Female	758.00	0.00	0.94	0.42	0.16	
Reading skill	Male	737.00	0.00	47.00	20.83	10.11	0.12
	Female	767.00	0.00	47.00	22.08	10.22	
State science test	Male	369.00	312.00	597.00	430.47	45.82	0.34
	Female	375.00	312.00	553.00	416.10	38.90	
Course grade	Male	772.00	–3.41	2.52	–0.11	1.02	0.23
	Female	826.00	–3.20	2.37	0.11	0.92	
Multiple-choice comprehension	Male	710.00	0.00	1.00	0.43	0.25	0.17
	Female	782.00	0.00	1.00	0.39	0.23	
Open-ended comprehension	Male	660.00	0.00	0.54	0.14	0.13	0.08
	Female	723.00	0.00	0.63	0.13	0.12	

Note. For each variable, there were significant effects of gender at the p < .02 level.

our analyses on the remaining 75% of the data (n = 1,229). As can be seen from the table, the students' scores increased as a function of grade level. This trend was confirmed by significant differences in grade level for strategy knowledge, $F(3, 1,167) = 3.18$, $p < .025$; science knowledge, $F(3, 1,163) = 7.54$; reading skill, $F(3, 1,223) = 22.19$; multiple-choice, $F(3, 1,170) = 8.65$; and open-ended comprehension questions, $F(3, 1,089) = 3.18$, $p < .05$. However, there were no significant differences in students' grade level on state science scores, $F(3, 504) = 1.90$, $p > .05$; and course grade, $F(3, 1,194) = 1.07$,

Table 4
Means and Standard Deviations of the Proportion Correct (for Reading Skill Number Correct) for Independent and Dependent Variables Used in This Study as a Function of Students' Grade Level

Measure	Grade	*n*	Min	Max	*M*	*SD*
Strategy knowledge	9	490	0.04	1.00	0.42	0.16
	10	383	0.12	0.84	0.43	0.15
	11	239	0.13	0.96	0.46	0.17
	12	59	0.16	0.76	0.45	0.15
Science knowledge	9	490	0.06	0.89	0.42	0.17
	10	386	0.06	0.94	0.43	0.17
	11	233	0.11	0.89	0.48	0.17
	12	58	0.11	0.83	0.45	0.16
Reading skill	9	517	0.00	47.00	19.17	9.73
	10	407	1.00	45.00	21.79	9.79
	11	242	3.00	47.00	25.38	10.39
	12	61	0.00	45.00	22.97	11.47
State science test	9	248	312.00	584.00	423.21	45.21
	10	158	346.00	553.00	424.93	41.21
	11	89	361.00	521.00	420.64	34.68
	12	13	312.00	444.00	396.54	37.64
Course grade	9	501	–3.20	2.52	0.00	0.98
	10	400	–2.49	2.37	0.09	0.91
	11	237	–3.15	2.22	0.03	0.99
	12	60	–2.44	1.69	–0.09	0.91
Multiple-choice comprehension	9	489	0.00	1.00	0.37	0.22
	10	392	0.00	1.00	0.43	0.25
	11	234	0.00	1.00	0.46	0.26
	12	59	0.00	1.00	0.44	0.24
Open-ended comprehension	9	442	0.00	0.50	0.11	0.11
	10	372	0.00	0.63	0.13	0.12
	11	223	0.00	0.54	0.18	0.14
	12	56	0.00	0.54	0.14	0.14

$p > .05$. The lack of difference is expected because there is no theoretical reason why state science scores and course grade should increase as a function of students' grade level. The result that on many of the measures students in the 11th-grade level scored numerically, but not statistically, higher than students in the 12th-grade level may be due to a confound between grade level and course difficulty. Whereas 62% of students in the 11th-grade level were enrolled in an advanced courses (e.g., honors, chemistry), only 41% of the 12th-grade-level students were enrolled in an advanced course. Accordingly, due to the significant amount of missing data, and the confound between grade level and course difficulty, we do not perform any additional analyses as a function of grade level.

Table 5
Correlations Between Science Achievement and Cognitive Abilities

	Multiple-Choice Comprehension	Open-Ended Comprehension	Course Grade	State ScienceTest
Science knowledge	r (1,389) = .505	r (,1290) = .549	r (1,438) = .245	r (685) = .594
Reading skill	r (1,442) = .527	r (1,345) = .641	r (1,456) = .244	r (692) = .582
Strategy knowledge	r (1,427) = .234	r (1,329) = .247	r (1,510) = .202	r (705) = .155

Note. All correlations were significant at the $p < .001$ level.

Table 6
Correlations Between Measures of Science Achievement

	Multiple-Choice Comprehension	Open-Ended Comprehension	Course Grade	State Science Test
Multiple-choice comprehension	—			
Open-ended comprehension	r (1,365) = .563	—		
Course grade	r (1,442) = .189	r (1,341) = .242	—	
State science test	r (691) = .562	r (640) = .548	r (712) = .331	—

Effect of Cognitive Abilities on Science Achievement

Our first question concerned the relative importance of the three cognitive measures in terms of predicting scores on each of the four measures of science achievement. Table 5 lists the individual correlations between the cognitive abilities and science achievement measures. All correlations were significant at the $p < .001$ level. As shown in Table 5, both science knowledge and reading skill had moderate to high correlations with the science achievement measures. Although the correlation values for metacognitive reading strategies were significant, the correlations were generally lower as compared to the correlations for reading skill and science knowledge. The correlation of metacognitive reading strategies to reading skill and to science knowledge were r (1,436) = .293 and r (1,452) = .271, respectively. The correlation between science knowledge and reading skill was r (1,433) = .577. Table 6 displays the correlations between the measures of science achievement. Science course grade exhibited the lowest correlations with the other three science achievement measures, multiple-choice, r (1,442) = .189; open-ended r (1,341) = .242; state science test, r (712) = .331.

Table 7 displays the correlations between the cognitive abilities and science achievement measures as a function of gender. All correlations were significant at the $p < .001$ level, with the exception of the correlation between strategy knowledge and state science scores for females, which was significant at $p = .001$. However, there were no significant differences between

Table 7
The Correlations Between the Ability and Achievement Measures as a Function of Gender

	Gender	Multiple-Choice Comprehension	Open-Ended Comprehension	Course Grade	State Science Test
Science knowledge	Males	$r(675) = .516$	$r(627) = .575$	$r(703) = .276$	$r(343) = .623$
	Females	$r(714) = .483$	$r(663) = .511$	$r(735) = .247$	$r(342) = .539$
Reading skill	Males	$r(698) = .573$	$r(651) = .656$	$r(711) = .231$	$r(347) = .604$
	Females	$r(744) = .497$	$r(694) = .642$	$r(745) = .250$	$r(345) = .596$
Strategy knowledge	Males	$r(682) = .259$	$r(636) = .319$	$r(735) = .216$	$r(351) = .180$
	Females	$r(745) = .232$	$r(693) = .196$	$r(775) = .171$	$r(354) = .177$

Note. All correlations were significant at $p < .001$, with the exception of the correlation between strategy knowledge and state science scores for females, which was significant at $p = .001$. There were no significant differences between males and females for any of the correlations at the $p < .001$ level.

Table 8
Regression Statistics, Beta Values, and Structure Coefficients for Each Measure of Science Achievement

Ability	Multiple-Choice Comprehension $F(4, 1{,}353) = 185.75$ $R^2 = .36$	Open-Ended Comprehension $F(4, 1{,}264) = 278.71$ $R^2 = .47$	Course Grade $F(4, 1{,}355) = 34.77$ $R^2 = .09$	State Science Test $F(4, 661) = 137.96$ $R^2 = .46$
Science knowledge	$t(1{,}353) = 9.35$ $\beta = .259$ $r_s = .847$	$t(1{,}264) = 9.23$ $\beta = .239$ $r_s = .798$	$t(1{,}355) = 5.07$ $\beta = .167$ $r_s = .773$	$t(661) = 9.63$ $\beta = .345$ $r_s = .866$
Reading skill	$t(1{,}353) = 13.74$ $\beta = .379$ $r_s = .906$	$t(1{,}264) = 19.46$ $\beta = .499$ $r_s = .940$	$t(1{,}355) = 3.10$ $\beta = .102$ $r_s = .767$	$t(661) = 11.15$ $\beta = .395$ $r_s = .861$
Strategy knowledge	$t(1{,}353) = 2.33$ $\beta = .054$ $r_s = .380$ $p = .02$	$t(1{,}264) = 2.14$ $\beta = .046$ $r_s = .362$ $p = .03$	$t(1{,}355) = 3.42$ $\beta = .094$ $r_s = .597$	n.s.
Gender (negative weights indicate males score higher)	$t(1{,}353) = -3.77$ $\beta = -.085$ $r_s = -.150$	$t(1{,}264) = -3.47$ $\beta = -.074$ $r_s = -.111$	$t(1{,}355) = 4.34$ $\beta = .116$ $r_s = .362$	$t(661) = -5.01$, $\beta = -.148$ $r_s = -.267$

Note. The three cognitive abilities and gender were entered in a regression equation to predict each achievement measure. All reported structure coefficients were significant at the $p < .001$ level.

males and females on any of the correlations at the $p < .001$ level. In general, the patterns were similar when the data were collapsed across gender.

To further investigate the relationships between the predictors and measures of science achievement, four sets of regressions, one for each science measure, were conducted. For each model, four predictor variables were entered: science knowledge, reading skill, reading strategy knowledge, and gender. Table 8 presents the regression statistics, beta values, and structure coefficients as a function of the cognitive abilities, gender, and the dependent science achievement measures.

For multiple-choice comprehension questions, the overall model accounted for 36% of the variance, $F(4, 1{,}353) = 185.75$; both reading skill and science knowledge reliably contributed to the model, whereas reading strategy knowledge was marginally significant ($p = .02$). Gender reliably contributed to the model; however, the beta weight was negative. A negative beta weight indicates that males scored higher on the entity in question or contributed positively to the model.

In terms of open-ended comprehension questions, the model accounted for 47% of the variance, $F(4, 1{,}264) = 278.71$. Similar to the results for the multiple-choice questions, both science knowledge and reading skill were significant predictors in the regression model for open-ended comprehension question scores, whereas reading strategy knowledge was only marginally significant ($p = .03$). Gender was also a significant predictor, indicating that males scored higher on the open-ended questions than females.

For students' course grade, the overall model accounted for only 9% of the variance, $F(4, 1{,}355) = 34.77$, and all three cognitive measures were significant predictors in the model. Gender was also significant, indicating that females scored higher than males on their science course grades.

For students' state science scores, the overall model accounted for 46% of the variance, $F(4, 661) = 137.96$. The students' reading skill and science knowledge were significant predictors in the model; however, the reading strategy knowledge score did not significantly contribute to the model. Gender was significant in the model, meaning that males scored higher than females on the state science test.

In summary, both reading skill and science knowledge were significant predictors (in the regression equations) for all of the science achievement measures. Reading skill had larger effects (i.e., beta weights) than science knowledge for scores on the passage comprehension questions (both multiple-choice and open-ended) but seemed to contribute similarly to course grade and state science scores. Reading strategy knowledge was a significant component of the regression equations for course grade and marginally significant for passage comprehension; however, strategy knowledge was not a significant contributor to state science scores. Gender was a significant predictor in all models, whereby males scored higher on the comprehension and state science scores and females scored higher on the final course grade.

Table 9

Statistics and Effect Sizes for Each Measure of Science Achievement, as a Function of Reading Skill and Science Knowledge

Ability	Multiple-Choice Comprehension	Open-Ended Comprehension	Course Grade	State Science Test
Science knowledge	$F(1,373) = 96.96$ Cohen's $d = 0.51$ $M_{lo} = 0.34$, $SD = 0.23$ $M_{hi} = 0.46$, $SD = 0.24$	$F(1,280) = 114.64$ Cohen's $d = 0.58$ $M_{lo} = 0.09$, $SD = 0.11$ $M_{hi} = 0.16$, $SD = 0.16$	$F(1,382) = 14.13$ Cohen's $d = 0.20$ $M_{lo} = -0.10$, $SD = 1.01$ $M_{hi} = 0.11$, $SD = 1.05$	$F(675) = 79.80$ Cohen's $d = 0.71$ $M_{lo} = 404.94$, $SD = 39.02$ $M_{hi} = 432.81$, $SD = 39.67$
Reading skill	$F(1,373) = 142.34$ Cohen's $d = 0.64$ $M_{lo} = 0.32$, $SD = 0.23$ $M_{hi} = 0.47$, $SD = 0.24$	$F(1,280) = 216.91$ Cohen's $d = 0.78$ $M_{lo} = 0.08$, $SD = 0.12$ $M_{hi} = 0.17$, $SD = 0.11$	$F(1,382) = 22.02$ Cohen's $d = 0.25$ $M_{lo} = -0.13$, $SD = 1.03$ $M_{hi} = 0.13$, $SD = 1.05$	$F(675) = 91.05$ Cohen's $d = 0.73$ $M_{lo} = 403.99$, $SD = 36.28$ $M_{hi} = 433.76$, $SD = 45.01$

Does Reading Skill Help Both Lower and Higher Science Knowledge Students?

The following analyses were conducted to examine whether reading skill improved achievement for both lower and higher knowledge students. One question of particular interest was whether reading skill helps lower knowledge students partially compensate for knowledge deficits—or, alternatively, whether science achievement is compromised regardless of reading skill for students without sufficient knowledge. To answer this question, ANOVAs were conducted including the between-subjects variables of science knowledge and reading skill.[1] Students were categorized as higher or lower science knowledge, and skilled or less skilled readers, by performing a median split on the respective measures. For each of the four measures of science achievement, an ANOVA was performed.

Figure 1 shows the effects of knowledge and reading skill on each of the four measures. As shown in Table 9, the main effects of science knowledge and reading skill were noteworthy for all four measures. There were also noteworthy interactions for multiple-choice questions, open-ended questions, and science course grade, but not state science scores, $F_{MCques}(1, 373) = 21.80$; $F_{OEques}(1, 280) = 7.74$; $F_{Grade}(1, 382) = 22.02$; $F_{State\ Science}(675) < 1$. The interactions reflected larger effects of reading skill for the higher knowledge participants. This is explored more fully below.

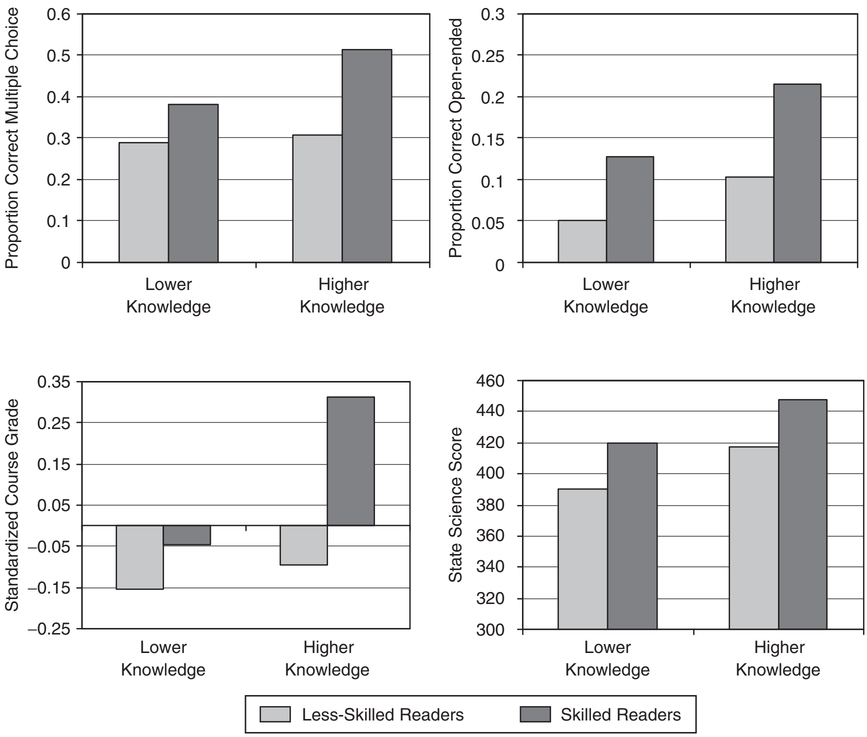

Figure 1. **Scores on the four science achievement measures as a function of science knowledge and reading skill.**

To answer our primary question, whether reading skill helps both higher and lower knowledge students, we performed separate tests to determine the effects of reading skill for both higher and lower knowledge readers. One central question was whether reading skill would partially compensate for knowledge. If reading skill helps compensate for science knowledge, then an effect of reading skill should be evident for lower knowledge students. Thus, the following *t* tests were conducted to determine if there were simple effects of reading skill for the lower knowledge participants. The effects of reading skill for the lower knowledge participants were noteworthy for three of the four measures, $t_{\text{MCques}}(651) = 5.65$, Cohen's $d = .46$; $t_{\text{OEques}}(579) = 11.12$, Cohen's $d = .89$; $t_{\text{State Science}}(250) = 7.15$, Cohen's $d = .96$. For course grade, reading skill showed a positive effect for lower knowledge readers, but the effect was quite small, $t(663) = 1.38$, $p = .17$; Cohen's $d = .12$. However, the effect of metacognitive reading strategy knowledge on course grade was small for lower knowledge readers, $t(666) = 2.19$, $p = .029$, Cohen's $d = .17$).

A more stringent test of compensation is to examine whether skilled, lower knowledge students score similarly to the less skilled, higher knowledge students. If the skilled, lower knowledge readers score as well as the less skilled, higher knowledge students, then reading skill can partially compensate for lower knowledge. Whereas the comparison of the means between the skilled, lower knowledge students and less skilled, higher knowledge students revealed that the skilled, lower knowledge students scored numerically higher than the less skilled, higher knowledge students, the differences were very small for measures of multiple-choice comprehension, $t(396) = 1.22$, $p > .05$, Cohen's $d = .14$ ($M_{\text{Less Skilled higher k}} = .35$, $SD = .21$; $M_{\text{Skilled lower k}} = .38$, $SD = .22$); state science scores, $t(190) < 1$, Cohen's $d = .06$ ($M_{\text{Less Skilled higher k}} = 417.56$, $SD = 35.29$; $M_{\text{Skilled lower k}} = 419.46$, $SD = 31.29$); and final course grade, $t(397) < 1$, Cohen's $d = .06$ ($M_{\text{Less Skilled higher k}} = -.10$, $SD = 1.02$; $M_{\text{Skilled lower k}} = -.04$, $SD = .95$). However, for open-ended passage comprehension, skilled, lower knowledge students scored higher than less skilled, higher knowledge readers, $t(371) = 2.56$, $p = .011$, Cohen's $d = .30$ ($M_{\text{Less Skilled higher k}} = .10$, $SD = .09$; $M_{\text{Skilled lower k}} = .13$, $SD = .11$).

We also performed the same set of ANOVAs mentioned above with the addition of gender as an independent variable. There were no three- or two-way interactions involving gender at the $p < .01$ level. Accordingly, the compensation effects described above apply equally for both males and females. In sum, the analyses reveal that skilled students were partially able to compensate for lower knowledge. Skilled, lower knowledge students scored as well or better than less skilled, higher knowledge students, indicating that reading skill can partially overcome knowledge deficits.

Second, we examined whether reading skill provided additional benefits to higher knowledge students. Individual *t* tests revealed the effect of reading skill was noteworthy for higher knowledge students for all four measures, $t_{\text{MCques}}(722) = 10.97$, Cohen's $d = .94$; $t_{\text{OEques}}(701) = 10.73$, Cohen's $d = 1.01$; $t_{\text{Grade}}(719) = 5.28$, Cohen's $d = .43$; $t_{\text{State Science}}(425) = 7.28$, Cohen's $d = .81$. These results held up at the $p < .001$ level for both males and females when analyzed separately. Thus, males and females both benefit from reading skill if they are higher knowledge students.

In summary, there were main effects of both reading skill and science knowledge for all four measures of science achievement. Although the effect of reading skill was greater for higher knowledge participants, reading skill had a noteworthy impact for lower knowledge students as well.

Discussion

The overarching aim of this study was to examine the impact of reading skill and science knowledge on high school students' content-based science achievement within classroom settings. We were particularly interested in examining the type of science achievement as measured by more standardized efforts such as the type of content-based assessment used in high-stakes testing. Whereas the traditional content-based approaches have been under attack in recent years (Carnoy, 2005; Chester, 2005a, 2005b; Chinn & Malhotra, 2002; Hess, 2005), these measures continue to influence policy.

This study has some notable strengths. A large sample of high school students was selected to obtain a range of students in terms of locality, SES, and ethnic makeup. Likewise, several different measures of content-based science achievement were utilized, including science passage comprehension (both multiple-choice and open-ended questions), students' course grade, and a statewide measure of science course knowledge (state science test). We also examined the relationships between the cognitive factors and science achievement as a function of gender.

There were three main research goals of the study: to determine the relative importance of reading abilities and knowledge in predicting our measures of science achievement and to determine whether reading skill compensated for knowledge deficits. We were also interested in assessing whether reading skill provided benefits above and beyond the impact of science knowledge; that is, whether reading skill reliably enhanced achievement for higher knowledge students. As a third aim, we were interested in examining whether there were any gender differences on our measures of cognitive abilities and science achievement.

In light of the first aim, we conducted both correlation and regression analyses to assess the relation between the cognitive factors and science achievement. The correlational analyses revealed that both measures of reading skill (Gates-MacGinitie and Metacomprehension Strategy Index) and science knowledge were moderately related to all measures of science achievement. A set of regression analyses revealed that scores on the science knowledge and reading skill measures were significant predictors (in a model that included science knowledge, reading skill, and strategy knowledge scores) for all of the dependent measures.

On the surface, it was surprising that the regression analyses revealed that cognitive abilities collectively predicted only 9% of the variance of the students' course grade and that there were small correlations with course grade and the other measures of achievement. However, course grade is a very complicated construct that is influenced by a large number of variables from simple class attendance to students' knowledge of the course material. Although speculative, one possibility is that some factors weighing on course grade may reflect teachers' orientation to use more deeper and authentic approaches to science instruction as opposed to more content and fact-based assessments. In turn, our individual measures of cognitive ability may not capture much of the deeper constructs that weigh in on course grades. However, this logic is speculative and future studies will be necessary to discern why our *individual* measures did not predict a large amount of the variance in course grade.

Interestingly, our results indicated that students who were *both* skilled and higher knowledge readers had much higher course grades than any of the three other conditions (less skilled, lower knowledge; skilled, lower knowledge; less skilled, higher knowledge). One possibility for this effect is that both knowledge *and* the ability to use knowledge (reading skill) are important for course grade. That is, neither ability by itself is sufficient to influence course grade, but both abilities interact such that higher course grades occur for the skilled, higher knowledge students. The interactivity between knowledge and reading skill may explain why the *individual* effects of the cognitive factors were weak predictors of course grade. In any event, it is clear that the measure of course grade is qualitatively different from the state science and passage comprehension measures. Indeed, state science and passage comprehension measures were more related to each other than with course grade.

We found that our measure of metacognitive reading strategy knowledge was a weak predictor of science achievement and significant only in the model predicting course grade. The design of our study does not allow us to determine precisely why metacognitive reading strategy knowledge only predicted course grade. However, we can speculate that the small relationship of metacognitive reading strategy knowledge and course grade may have occurred because our measure of metacognitive reading strategy knowledge may uniquely tap into processes related to course success such as planning, efficient use of time, and study strategies. Although the correlations with the measure of metacognitive reading strategy knowledge were lower than the correlations for the other cognitive variables, this does not necessarily mean that metacognition is not as important as the other variables. One possible explanation for the weak relationship could be due to the sensitivity of our measure, and as a result our findings may underestimate the impact of metacognitive knowledge. Although there have been several efforts to create highly sensitive and valid measures of metacognitive ability, many assessments are lacking, and there is strong need to construct better measures (Mokhtari & Reichard, 2002). Future work is needed to develop and test more effective assessments of metacognition and to test these measures in large-scale research settings.

In light of the second aim, we categorized participants as higher and lower science knowledge and skilled and less skilled readers to determine whether reading skill offsets students' knowledge deficits or, conversely, only benefited higher knowledge students. For lower knowledge students, reading skill significantly improved comprehension for three out of the four science achievement measures. For the one measure in which reading skill did not significantly help lower knowledge readers (i.e., course grade), a high level of reading strategy knowledge did improve achievement. In fact, reading skill helped lower knowledge students score as high or higher than less skilled, higher knowledge readers. Thus, with our measures of achievement, reading skill helped compensate for low knowledge. In addition, for higher knowledge students, a high level of reading skill significantly enhanced achievement on all four measures beyond the effects of knowledge. In particular, our data suggest that having both a higher level of knowledge and reading skill is the best situation.

The robust ability of the cognitive factors to predict science achievement is not surprising given that laboratory research has previously established the influence of reading skill, reading strategies, and knowledge on content learning (Chi et al., 1989; Dochy et al., 1999; Voss & Silfies, 1996). However, the current work builds upon previous research by demonstrating the important roles of reading strategy knowledge, reading skill, and science knowledge for achievement in real-world high school classes and, in particular, state-mandated science achievement measures. Field studies are essential because research has shown that the conclusions drawn from laboratory studies may differ dramatically from the conclusions taken from similar research in more ecologically valid settings (Ceci & Bronfenbrenner, 1985; DeLoache, Cassidy, & Brown, 1985). The present results demonstrate that the laboratory findings on the impact of science knowledge, reading skill, and metacognitive reading strategies extend to classroom environments.

An important goal of the present work was to discern whether metacognitive reading strategies or general reading skill could combat the knowledge deficits faced by high school students. Such results may point toward the potential success of reading interventions on content-based science achievement, particularly those that teach students to use reading strategies. Our results were indeed hopeful in that regard. In general, the data support the notion that reading skill can help lower knowledge learners improve their content-based science achievement scores. More skilled readers consistently scored higher than less skilled counterparts on all of our measures of science achievement.

One way to examine the compensatory effects of reading skill for knowledge deficits is to examine whether the effects are significant for lower knowledge students, which we found to be the case. Another is to look at whether reading skill brings the lower knowledge student up to par with the higher knowledge student. In all cases (see Figure 1), the skilled, lower knowledge students performed as well as the less skilled, higher knowledge students. Thus, whereas having both abilities is obviously ideal for success in science, one or the other can contribute to better science understanding. In short, our results are consistent with Adams et al. (1995) in that we also found that reading skill can compensate for knowledge deficits.

These results are particularly important because research has shown that students are often not equipped with the knowledge necessary to succeed in their courses (Snow, 2002), and high school texts are often incomplete because they fail to provide necessary background information (Beck et al., 1989). Conversely, the results of this study suggest that reading skill (and to a lesser extent metacognitive reading strategies) can at least partially overcome content-based knowledge deficits.

As our third aim, we explored whether there were any gender differences on our measures of cognitive ability and science achievement. We found that males scored higher than females on measures of science knowledge, state science test, and passage comprehension. These results are consistent with research that indicates that males are more likely to score higher than females on measures of science content (Kahle, 2004; Mau & Cheng, 2000; Reis & Park, 2001). In contrast, females scored higher than males on measures of reading skill, strategy knowledge, and course grade. These findings are consistent with research that has shown that females score higher than males on measures of reading comprehension (Mau & Cheng, 2000) and females often use more study strategies than males (Slotte et al., 2001). The result that females in this study had higher course grades is consistent with research that found females had higher GPAs than males (Rech, 1996). Whereas gender differences were evident in this study, it is important to keep in mind that the effect sizes were relatively small (average effect size Cohen's $d = .19$) and also that gender differences are becoming smaller in recent years (Fergusson & Horwood, 1997; Myers, 2002).

Limitations

Although the current study has some clear strengths, there are several limitations. First, the variables used in this study were quasi in nature: We did not directly manipulate science knowledge, reading skill, or metacognitive reading strategy knowledge. Thus, we can only make claims about the possible relationships between the variables; we cannot make any claims about causation. Second, we only had access to the state science tests for two of the four schools in this study. Third, we separated participants into groups of *higher* and *lower* science knowledge based on our sample. This constitutes a *relative* measure of higher and lower knowledge that is dependent upon the students included in our sample. Our measure does not mean to imply an absolute measure of high and low knowledge students as compared to the population as a whole.

Fourth, in our design we presented the materials in a fixed order (i.e., science passage, then reading skill test, then science knowledge test, etc.). One might argue that the order of these materials should have been counterbalanced across participants to avoid any potential negative effects of task order on the results. However, this was a well-considered design choice on our part. We presented the tests in this manner, rather than counterbalancing the tests, for two reasons. First, we needed to control for potential effects of the science knowledge test priming concepts in the comprehension test. This can have differential effects as a function of knowledge. For example, McNamara and Kintsch (1996) gave half of the participants a knowledge test before reading the text. They found that higher knowledge participants who received the knowledge test before reading the text performed better on the comprehension test for the text than higher knowledge participants who did not receive a knowledge test before reading the text, presumably because the knowledge test primed certain concepts (even though it was carefully controlled such that there was no overlap in information). In contrast, lower knowledge participants performed worse than their counterparts if they received the prior knowledge test before reading the text. Here our aim was to avoid such a situation; hence, the knowledge test was presented after the comprehension text and test. Although, conversely, the comprehension text and test may prime concepts in the science knowledge test, it would have had an equal effect for all participants, and we avoided the situation where half of the participants would have had one situation and half another. We also avoided the situation of exacerbating knowledge differences on the comprehension test—a primary measure of interest.

A second reason for not counterbalancing the tests concerns fatigue effects. Much as in a standardized test situation, presenting tasks in a fixed order controls fatigue effects on each task so that all participants have experienced relatively equal fatigue before each task. This avoids the situation where half of the participants are fatigued when they are presented one of the tasks such as the text and questions, and half are not. Essentially, tests such as the ones presented in this experiment are often not counterbalanced so that comparisons between participants on the tests control for fatigue and practice effects—essentially, it is often considered better if all of the participants are equally fatigued for any particular test. A drawback of this approach is of course that there are potential resulting task order effects.

Finally, although not a limitation, it is important to remember that our study was designed to examine the impact of content-based measures of science achievement, such as those typically found on high-stakes tests. Consequently, the conclusions based on our findings may not generalize to more authentic approaches to science and inquiry learning.

Suggestions for Future Research

So what can a student (or teacher) do to improve achievement on content-based measures of science? The results of this study suggest several areas of further research. First, the findings support the notion that reading skill is important for content-based science achievement. In fact, reading skill was one of the best single predictors of student achievement. Reading skill was also positively related to metacognitive reading strategies, science knowledge, and all of the measures of content-based achievement. More important, the results indicate that increased reading skill can help learners to partially compensate for lower knowledge: Skilled readers performed better on measures of science achievement even when knowledge was impoverished. This is especially important when students are faced with situations in which their knowledge base is poor, for instance, when they read difficult texts and when they begin new courses. Consequently, future research should look at how teachers and parents can foster positive attitudes toward reading and encourage students to read more. Research has shown that an increase in exposure to print is associated with a corresponding increase in reading skill (Cipielewski & Stanovich, 1992). Our findings may suggest that further research is needed to determine whether programs aimed at improving reading skill can also help improve students' content-based science achievement.

Second, the results underscore the importance of science knowledge on achievement: Science knowledge was highly correlated with all measures of science achievement. Research has shown that knowledge is critical in determining how well new information is learned (Dochy et al., 1999; Shapiro, 2004; Spilich et al., 1979; Willoughby et al., 1993, 1994). In terms of application, we suggest that students should gain more knowledge by reading books, reading science articles, and watching educational programs. The more knowledge students gain, the easier it should be to learn new information. However, one of the biggest impediments for problem-solving success is failure to retrieve the appropriate knowledge and *apply* knowledge in a suitable manner (Reeves & Weisberg, 1994). Unfortunately, many students either fail to draw on appropriate knowledge (Bransford & Johnson, 1972) or fall short of making complete use of their knowledge (Wood et al., 1990). Surprisingly, even gifted learners sometimes do not utilize their background knowledge to learn new material (Wood & Hewitt, 1993). The challenge of educators is to discover ways not only to improve the acquisition of new information (i.e., knowledge) but also to ensure that students access and apply that information appropriately. Further research is needed to determine the precise impact of how knowledge helps students learn new information.

However, as we well know, students will often find themselves in courses for which they are ill prepared. In that case, knowing and using metacognitive reading strategies may help the learner to partially overcome knowledge deficits. Unfortunately, students do not automatically engage in strategic processing (Garner, 1990; Pressley et al., 1992; Pressley & Ghatala, 1990), especially students who have comprehension difficulties (Lenski & Nierstheimer, 2002). Self-regulated learning occurs much less frequently than believed, and even college students often fail to monitor their learning (Pressley & Ghatala, 1990). To make matters worse, although the benefits of metacognition on learning are clear, metacognitive strategies are not being used in the classrooms as they should be (Cox, 1997). Consequently, more research should be conducted to discover whether techniques that promote metacognitive strategy use help content-based learning.

Investigations on this matter have proved successful. For example, McNamara (McNamara, 2004; McNamara & Scott, 1999) developed a reading strategy program called Self Explanation Reading Training (SERT).[2] The SERT strategy is designed to teach students to promote higher level reading skills and metacognitive ability by making use of predictions, elaborations, bridging inferences, paraphrasing, comprehension monitoring, and logic. The results with college students (McNamara, 2004) and high school students (O'Reilly et al., 2004) have been encouraging: Participants trained to use the SERT strategy outperformed control participants on measures of comprehension with difficult science passages. More important, the benefits of SERT training occurred for students with lower knowledge, that is, for those that needed the most help. However, it is important to point out that although reading strategy interventions have been shown to be successful in improving content-based achievement, to date we are not aware of any study that has examined the impact of strategy training on more deeper and authentic measures of science achievement. Future studies and interventions need to be developed to determine the long-term benefits of strategy training on more authentic or process measures of science achievement.

Finally, we also recommend addressing gender differences by working together with teachers and parents to reduce stereotypical attitudes and behaviors that support gender differences. This is important because research has shown that gender differences are supported by socialization. For example, recent work has shown that mothers engage in more science talk with boys than they do with girls (Tenenbaum, Snow, Roach, & Kurland, 2005), and teachers are often more responsive with boys than girls (She, 2001). We urge educators and parents to be attentive to these biases.

We are mindful that the present findings are based on quasi-experimental variables and should be interpreted with caution. Despite this limitation, the conclusion we draw from this work is that cognitive abilities are important for science achievement as measured by the types of items typically included on high-stakes tests and that reading skill can partially compensate for knowledge deficits. This study also builds on previous research by demonstrating the effects of cognitive abilities on content-based science achievement in a large school setting.

APPENDIX

Science Knowledge Test

1. For any laboratory experiment, what should be the first step?
   a. Form a hypothesis.
   b. State the problem.
   c. Perform the experiment.
   d. Write the conclusion.

2. When measuring a small quantity of liquid, which instrument would permit the most accurate reading?
   a. beaker
   b. flask
   c. test tube
   d. graduated cylinder

3. Today, Marjorie's lab work involves working with bases. Which types of safety equipment is Marjorie most likely to need?
   a. gloves and safety glasses
   b. safety glasses and litmus paper
   c. a fire extinguisher and a blanket
   d. litmus paper and an apron

4. The New Wave Swim Team uses a 50-meter pool during the summer to prepare for long-course events. Josh competes in the 1,500 meter. How many kilometers are in this event?
   a. 1.5
   b. 15
   c. 50
   d. 150

5. The shape of the Earth is________
   a. round.
   b. obtuse.

c. spherical.
d. circular.

Read the passage and answer the three questions that follow.

> A tropical fish hobbyist wants to see whether a new, more expensive fish food causes fish to grow faster. She sets up an experiment using four groups of ten fish each. The fish in Group A receive a diet of the new fish food. Fish in Group B receive the old fish food. Fish in Group C receive a mixture containing two parts new food and one part old food. Fish in Group D receive a mixture containing three parts new food and one part old food. The length of each fish is measured weekly.

6. What is the control group in this experiment?
   a. Group A
   b. Group B
   c. Group C
   d. Group D

7. What is the independent variable in this experiment?
   a. the fish
   b. the type of fish food
   c. the length of the fish
   d. expensive fish food

8. What is the dependent variable in this experiment?
   a. the fish
   b. the type of fish food
   c. the length of the fish
   d. expensive fish food

9. Which unit of measurement is largest?
   a. yard
   b. foot
   c. meter
   d. millimeter

10. A man shopping for a car wants to calculate the average price for a car at seven different dealers. How can he do this?
    a. Add the numbers together.
    b. Identify the most frequent number.
    c. Sort the numbers and identify the fourth lowest number.
    d. Add the numbers together and divide by seven.

11. Which of the following numbers represents the ratio of 10 to 1?
    a. 10/1
    b. 1/10
    c. 1*10
    d. 10-1

12. What fraction of a yard is 2 feet?
    a. 1/10
    b. 2/3
    c. 3/4
    d. 1/3

13. The lowest level of the earth's atmosphere is the ___________
    a. stratosphere.
    b. troposhere.
    c. air.
    d. ozone.

14. An inversion refers to ______________
    a. a sudden change in weather conditions.
    b. warmer air near the earth's surface than the air above it.
    c. pollution and smog.
    d. colder air near the earth's surface than the air above it.

15. Dark, gray layers of cloud that produce steady rain are called ___________
    1. cumulus clouds.
    2. cirrus clouds.
    3. nimbostratus clouds.
    4. cirrocumulus clouds.

16. The change from water vapor to liquid water is called ___________
    a. condensation.
    b. evaporation.
    c. humidity.
    d. liquefaction.

17. When warm air replaces an equal volume of cold air,
    a. the humidity increases.
    b. the air pressure at the ground decreases.
    c. the air pressure at the ground increases.
    d. an inversion occurs.

18. High air pressure indicates that the air is ___________
    a. dense.
    b. light.
    c. humid.
    d. at high elevations.

Notes

We would like to thank the members of the ODU Strategy Lab who helped to conduct this study, including Kim Cottrell, Meghan Depont, Karen Fuller, Erin McSherry, Raymond Morgan, Grant Sinclair, Danny Simmons, Karen Stockstill, and Paul Walker. This project was supported by an NSF IERI grant (REC-0241144) and an IES grant (R305G040046) to the second author. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or IES.

[1]We also conducted these analyses including reading strategy knowledge rather than reading skill. The trends were the same for all four measures. Hence, we are only presenting the results as a function of reading skill. Both reading skill and strategy knowledge could not be included because there were resulting cells with too few observations. A separate analysis that utilized a combined measure of reading skill and the strategy knowledge (both scores were converted to standardized scores and then averaged) revealed similar trends.

[2]An automated version of Self Explanation Reading Training (SERT) training called iSTART has also resulted in improved science comprehension for college and high school students (see McNamara, Levinstein, & Boonthum, 2004; O'Reilly, Sinclair, & McNamara, 2004).

References

Adams, B., Bell, L., & Perfetti, C. (1995). A trading relationship between reading skill and domain knowledge in children's text comprehension. *Discourse Processes, 20,* 307–323.

Alcock, K., Nokes, K., Ngowi, F., Musabi, C., Mbise, A., Mandali, R., et al. (2000). The development of reading tests for use in a regularly spelled language. *Applied-Psycholinguistics, 21,* 525–555.

Alexander, P. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. *Advances in Motivation and Achievement, 10,* 213–250.

Alexander, P., & Judy, J. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58,* 375–404.

Beck, I., McKeown, M., & Gromoll, E. (1989). Learning from social studies texts. *Cognition and Instruction, 6,* 99–158.

Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction, 2,* 131–156.

Bransford, J., Brown, A., & Cocking, R. (Eds.). (1999). *How people learn: Brain, mind, experience and school*. Washington, DC: National Academy Press.

Bransford, J., & Johnson, M. K. (1972). Contextual prerequisites for understanding some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

Bransford, J., Stein, B., Vye, N., Franks, J., Auble, P., Mezynski, K., et al. (1982). Differences in approaches to learning: An overview. *Journal of Experimental Psychology: General, 111,* 390–398.

Brown, A., Bransford, J., Ferrara, R., & Campione, J. (1983). Learning, remembering and understanding. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology* (Vol. 3, pp. 77–166). New York: John Wiley.

Carnoy, M. (2005). Have state accountability and high-stakes tests influenced student progression rates in high school? *Educational-Measurement: Issues and Practice, 24,* 19–31.

Ceci, S., & Bronfenbrenner, U. (1985). "Don't forget to take the cupcakes out of the oven": Prospective memory, strategic time-monitoring, and context. *Child Development, 56,* 152–164.

Chester, M. (2005a). Making valid and consistent inferences about school effectiveness from multiple measures. *Educational Measurement: Issues and Practice, 24,* 40–52.

Chester, M. (2005b). Measuring the impact of state accountability programs. *Educational Measurement: Issues and Practice, 24,* 3–4.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182.

Chi, M. T. H., De Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Chiesi, H. I., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 275–290.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic reasoning in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86,* 175–218.

Cipielewski, J., & Stanovich, K. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology, 54,* 74–89.

Cottrell, K., & McNamara, D. S. (2002). Cognitive precursors to science comprehension. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual*

Meeting of the Cognitive Science Society (pp. 244–249). Mahwah, NJ: Lawrence Erlbaum.
Cox, B. (1997). The rediscovery of the active learner in adaptive contexts: A developmental-historical analysis of transfer of training. *Educational Psychologist, 32,* 41–55.
DeLoache, J., Cassidy, D., & Brown, A. (1985). Precursors of mnemonic strategies in very young children's memory. *Child Development, 56,* 125–137.
Dochy, F., Segers, M., & Buehl, M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69,* 145–186.
Driscoll, M., Moallem, M., Dick, W., & Kirby, E. (1994). How does the textbook contribute to learning in a middle school science class? *Contemporary Educational Psychology, 19,* 79–100.
Ericsson, A., & Kintsch, W. (1995). Long term working memory. *Psychological Review, 102,* 211–245.
Fergusson, D., & Horwood, J. (1997). Gender differences in educational achievement in a New Zealand birth cohort. *New Zealand Journal of Educational Studies, 32,* 83–96.
Flannery, K. A., Liederman, J., Daly, L., & Schultz, J. (2000). Male prevalence for reading disability is found in a large sample of Black and White children free from ascertainment bias. *Journal of the International Neuropsychological Society, 6,* 433–442.
Forget, M. A. (1999). *Comparative effects of three methods of staff development in content area reading instruction on urban high school teachers*. Unpublished doctoral dissertation, Old Dominion University, Norfolk, VA.
Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Psychology, 60,* 517–529.
Graesser, A., Person, N., & Harter, D. (2001). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education, 12,* 23–39.
Graesser, A. C., & Wisher, R. A. (2001). *Question generation as a learning multiplier in distributed-learning environments* (Technical Report 1121). Alexandria, VA; U.S. Army Research Institute for the Social and Behavioral Sciences.
Hess, F. (2005). Commentary: Accountability policy and scholarly research. *Educational Measurement: Issues and Practice, 24,* 53–57.
Kahle, J. (2004). Will girls be left behind? Gender differences and accountability. *Journal-of-Research in Science Teaching, 41,* 961–969.
Kane, M. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice, 21,* 31–41.
Kesidou, S., & Roseman, J. E. (2002). How well do middle school science programs measure up? Findings from Project 2061's curriculum review. *Journal of Research in Science Teaching, 39,* 522–549.
Lane, S. (2004). Validity of high-stakes assessment: Are students engaged in complex thinking? *Educational Measurement: Issues and Practice, 23,* 6–14.
Langer, E. (1989). *Mindfulness.* Reading, MA: Addison-Wesley.
Lenski, S., & Nierstheimer, S. (2002). Strategy instruction from a sociocognitive perspective. *Reading Psychology, 23,* 127–143.
Liederman, J., Kantrowitz, L., & Flannery, K. (2005). Male vulnerability to reading disability is not likely to be a myth: A call for new data source. *Journal of Learning Disabilities, 38,* 109–129.
Long, D., & Golding, J. (1993). Superordinate goal inferences: Are they automatically generated during comprehension? *Discourse Processes, 16,* 55–74.
Long, D., Oppy, B., & Seely, M. (1994). Individual differences in the time course of inferential processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20,* 1456–1470.

Lord, F., & Novick, M. R. (1968). *Statistical theory of mental tests*. New York: Addison-Wesley.

Lumpe, A., & Beck, J. (1996). A profile of high school biology textbooks with implications for biology instruction. *American Biology Instructor, 58,* 147–153.

Lundeberg, M. (1987). Metacognitive aspects of reading comprehension: Studying understanding in legal case analysis. *Reading Research Quarterly, 22,* 407–432.

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

MacGinitie, W. H., & MacGinitie, R. K. (1989). *Gates-MacGinitie Reading Test* (3rd ed.). Chicago: Riverside.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure. *Cognition and Instruction, 21,* 251–283.

Magliano, J. P., Wiemer-Hastings, K., Millis, K. K., Muñoz, B. D., & McNamara, D. S. (2002). Using latent semantic analysis to assess reader strategies. *Behavior Research Methods, Instruments, and Computers, 34,* 181–188.

Mau, W., & Cheng, L. (2000). Gender differences in homework and test scores in mathematics, reading and science at tenth and twelfth grade. *Psychology, Evolution and Gender, 2,* 119–125.

Mayer, R. (2001). *Multimedia learning*. New York: Cambridge University Press.

McNamara, D., Kintsch, E., Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1–43.

McNamara, D. S. (1997). Comprehension skill: A knowledge-based account. In *Proceedings of Nineteenth Annual Meeting of the Cognitive Science Society* (pp. 508–513). Hillsdale, NJ: Lawrence Erlbaum.

McNamara, D.S. (2001). Reading both high and low coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology, 55,* 51–62.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38,* 1–30.

McNamara, D. S., de Vega, M., & O'Reilly, T. (in press). Comprehension skill, inference making, and the role of knowledge. In F. Schmalhofer & C.A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes*. Mahwah, NJ: Lawrence Erlbaum.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes, 22,* 247–288.

McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy trainer for active reading and thinking. *Behavioral Research Methods, Instruments, and Computers, 36,* 222–233.

McNamara, D. S., & McDaniel, M. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,* 465–482.

McNamara, D. S., & O'Reilly, T. (in press). Theories of comprehension skill: Knowledge and strategies versus capacity and suppression. In F. Columbus (Ed.), *Progress in experimental psychology research*. Hauppauge, NY: Nova Science.

McNamara, D. S., & Scott, J. L. (1999). Training reading strategies. In *Proceedings of the Twenty-First Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition, 29,* 10–17.

Mokhtari, K., & Reichard, C. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94,* 249–259.

Myers, D. G. (2002). *Social psychology* (7th ed.). New York: McGraw-Hill.

National Education Association. (2002). *2002 instructional materials survey: Report of findings September 2002*. Washington, DC: Author.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

Oakhill, J. (1983). Instantiation in skilled and less skilled comprehenders. *Quarterly Journal of Experimental Psychology, 35,* 441–450.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54,* 31–39.

Oakhill, J., & Yuill, N. (1996). Higher order factors in comprehension disability: Processes and remediation. In C. Cornaldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and Intervention*. Mahwah, NJ: Lawrence Erlbaum.

O'Reilly, T., Best, R., & McNamara, D. S. (2004). Self-explanation reading training: Effects for low-knowledge readers. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 1053–1058). Mahwah, NJ: Lawrence Erlbaum.

O'Reilly, T., & McNamara, D. S. (2002a, November). *Text coherence effects: Interactions of prior knowledge and reading skill*. Poster presented at the 24th annual meeting of the Psychonomic Society, Kansas, MO.

O'Reilly, T., & McNamara, D. S. (2002b). What's a science student to do? In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual Meeting of the Cognitive Science Society* (pp. 726–731). Mahwah, NJ: Lawrence Erlbaum.

O'Reilly, T. P., Sinclair, G. P., & McNamara, D. S. (2004). Reading strategy training: Automated versus live. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 1059–1064). Mahwah, NJ: Lawrence Erlbaum.

Ozuru, Y., Best, R., & McNamara, D. S. (2004). Contribution of reading skill to learning from expository texts. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 1071–1076). Mahwah, NJ: Lawrence Erlbaum.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction, 2,* 117–175.

Paris, S., & Jacobs, J. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development, 55,* 2083–2093.

Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76,* 1239–1252.

Perfetti, C. (1989). There are generalized abilities and one of them is reading. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 307–336). Hillsdale, NJ: Lawrence Erlbaum.

Perie, M., Grigg, W. S., & Donahue, P. L. (2005). *The nation's report card: Reading 2005* (NCES 2006-451; U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics). Washington, DC: Government Printing Office.

Pressley, M. (2002). Comprehension strategies instruction: A turn-of-the-century report. In C. C. Block & M. Pressley (Eds.). *Comprehension instruction: Research-based best practices* (pp. 11–27). New York: Guilford.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum.

Pressley, M., & Ghatala, E. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist, 25,* 19–33.

Pressley, M., Wood, E., Woloshyn, V., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychology, 27,* 91–109.

Radcliffe, R., Caverly, D., Peterson, C., & Emmons, M. (2004). Improving textbook reading in a middle school science classroom. *Reading Improvement, 41,* 145–157.

Rech, J. (1996). Gender differences in mathematics achievement and other variables among university students. *Journal of Research and Development in Education, 29,* 73–76.

Reeves, L., & Weisberg, R. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin, 115,* 381–400.

Reis, S., & Park, S. (2001). Gender differences in high-achieving students in math and science. *Journal for the Education of the Gifted, 25,* 52–73.

Roseman, J. E., Kesidou, S., Stern, L., & Caldwell, A. (1999). Heavy books light on learning: AAAS Project 2061 evaluates middle grades science textbooks. *Science Books & Films, 35,* 243–247.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64,* 479–530.

Ross, B. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning Memory, and Cognition, 15,* 456–468.

Ross, B., & Bradshaw, G. (1994). Encoding effects of remindings. *Memory & Cognition, 22,* 591–605.

Ross, B., & Kennedy, P. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning Memory, and Cognition, 16,* 42–55.

Saloman, G., & Perkins, D. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24,* 113–139.

Schmitt, M. (1990). A questionnaire to measure children's awareness of strategic reading process. *The Reading Teacher, 49,* 454–461.

Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal, 41,* 159–189.

She, H. (2001). Different gender students' participation in the high- and low-achieving middle school questioning-orientated biology classrooms in Taiwan. *Research in Science and Technological Education, 19,* 147–158.

Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. (1992). Individual differences in bridging inference processes. *Memory & Cognition, 20,* 539–548.

Singer, M., & Ritchot, K. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory & Cognition, 24,* 733–743.

Slotte, V., Lonka, K., & Lindblom, Y. (2001). Study-strategy use in learning from text. Does gender make any difference? *Instructional Science, 29,* 255–272.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Snow, C., Burns, M., & Griffin, P. (1998). *Preventing reading difficulties in children.* Washington DC: National Academy Press.

Spaulding, N., & Namowitz, S. (1994). *Heath earth science*. Washington, DC: Heath.

Spilich, G., Vesonder, G., Chiesi, H., & Voss, J. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 275–290.

Taylor, B. M., Pearson, P. D., Clark, K. F., & Walpole, S. (1999). Effective schools/accomplished teachers. *Reading Teacher, 53,* 156–159.

Tenenbaum, H., Snow, C., Roach, K., & Kurland, B. (2005). Talking and reading science: Longitudinal data on sex differences in mother-child conversations in low-income families. *Journal of Applied Developmental Psychology, 26,* 1–19.

Thompson, R., & Zamboanga, B. (2004). Academic apititude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology, 96,* 778–784.

Timmerman, B. (2002). Keeping science current. *Science Scope, 25,* 12–15.

Tobias, S. (1994). Interest prior knowledge and learning. *Review of Educational Research, 64,* 37–54.

VanLehn, K. (1990). *Mind bugs: The origins of procedural misconceptions*. Cambridge, MA: MIT Press.

VanLehn, K. (1995). *Rule learning events in the acquisition of a complete skill* (Technical Report of the Learning Research Development Center ONR-92-4). Pittsburgh, PA: University of Pittsburgh.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47,* 513–539.

VanLehn, K. (1998). Analogy events: How examples are used during problem solving. *Cognitive Science, 22,* 347–388.

Voss, J., & Silfies, L. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction, 14,* 45–68.

Walczyk, J., Marsiglia, C., Bryan, K., & Naquin, P. (2001). Overcoming inefficient reading skills. *Journal of Educational Psychology, 93,* 750–757.

Walker, C. (1987). Relative importance of domain knowledge and overall aptitude on acquisition of domain-related information. *Cognition and Instruction, 4,* 25–42.

Weiss, I., Pasley, J., Smith, P., Banilower, E., & Heck, D. (2003). *A study of K–12 mathematics and science education in the United States*. Chapel Hill, NC: Horizon Research.

Whitney, P., Ritchie, B., & Clark, M. (1991). Working memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes, 14,* 133–145.

Willoughby, T., Waller, T. G., Wood, E., & MacKinnon, G. E. (1993). The effect of prior knowledge on an immediate and delayed associative learning task following elaborative interrogation. *Contemporary Educational Psychology, 18,* 36–46.

Willoughby, T., Wood, E., & Khan, M. (1994). Isolating variables that impact or detract from the effectiveness of elaboration strategies. *Journal of Educational Psychology, 86,* 279–289.

Wilson, P. T., & Anderson, R. C. (1986). What they don't know will hurt them: The role of prior knowledge in comprehension. In J. Orasanu (Ed.), *Reading comprehension: From research to practice* (pp. 31–48). Hillsdale, NJ: Lawrence Erlbaum.

Woloshyn, V. E., Schnieder, W., & Pressely, M. (1992). Elaborative interrogation and prior-knowledge effects on learning facts. *Journal of Educational Psychology, 84,* 115–124.

Wood, E., & Hewitt, K. L. (1993). Assessing the impact of elaborative strategy instruction relative to spontaneous strategy use in high achievers. *Exceptionality, 4,* 65–79.

Wood, E., Pressley, M., & Winne, P. H. (1990). Elaborative interrogation effects on children's learning of factual content. *Journal of Educational Psychology, 82,* 741–748.

Yuill, N., & Oakhill, J. (1988). Understanding of anaphoric relations in skilled and less skilled comprehenders. *British Journal of Psychology, 79,* 173–186.

Zhang, K., & Zhang, B. (1996). Impact of interest on text comprehension. *Acta Psychologica Sinica, 28,* 284–289.

Manuscript received October 12, 2004
Revision received June 19, 2006
Accepted August 1, 2006