Applied Cognitive Psychology, *Appl. Cognit. Psychol.* **26**: 104–117 (2012)
Published online 24 May 2011 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/acp.1802

What Variables Predict Quality of Text Notes and are Text Notes Related to Performance on Different Types of Tests?

STEPHEN T. PEVERLY[1]* and JAMES F. SUMOWSKI[2]

[1]*Teachers College, Columbia University, New York, USA*
[2]*Kessler Foundation Research Center, West Orange, USA*

Summary: Despite the importance of notes to test performance, very little is known about the cognitive variables related to notetaking, especially text notes. The primary purpose of this study was to evaluate the contributions of transcription fluency, reading comprehension, verbal working memory, executive attention, and background knowledge to the quality of text notes. A secondary purpose was to evaluate the relationship of all of the other aforementioned variables to three test types to determine if notes are more important to some test types than others: a free recall essay and two types of multiple choice items: memory items and inference items. Results indicated that transcription fluency was the best predictor of notes (reading comprehension was also significant predictor), which extends previous findings on the importance of transcription fluency to lecture notes. Notes' were the best predictor of the essay and the memory multiple choice items but not the inference items. Copyright © 2011 John Wiley & Sons, Ltd.

Theories of performance in academic skills such as reading (Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; Vellutino, Fletcher, Snowling, & Scanlon, 2004), writing (Berninger et al., 2006; McCutchen, 1996), and mathematics (Geary, 1994) recognize that: (i) competence in academic skills requires the parallel execution of domain specific basic and higher level processes within a limited capacity working memory; (ii) domain-specific basic skills must be executed automatically/fluently so that limited working memory resources can be directed toward the application of higher level cognitive skills; and (iii) once basic skills are sufficiently automatic or fluent, competence in a skill is largely determined by the quality of higher level skills. In skilled reading, for example, Perfetti's (1985, 2007) Verbal Efficiency Theory suggests that comprehension requires the parallel execution of word recognition and the higher level cognitive processes needed to interpret text, such as language ability. Because comprehension occurs in a limited capacity working memory, word recognition must be automatized if language ability and the other higher level cognitive processes related to comprehension are to be applied effectively. Once word recognition is automatized, comprehension skill is most strongly related to readers' higher level cognitive processes.

Most research on the cognitive underpinnings of academic skills has focused on those that begin to develop in elementary school or earlier: reading, mathematics, and writing. There is comparatively little research on the cognitive variables that underlie academic skills that begin to develop in late middle school and high school. One particularly important late developing skill is notetaking. After elementary school, the amount of information students is required to understand increases dramatically (Thomas, Iventosch, & Rohwer, 1987) and most of the information is presented in forms students find difficult to process: lecture (Piolat, Olive, & Kellogg, 2005) and expository texts (Mulcahy-Ernt, & Caverly, 2009; Snow, 2002). Research, most of which has been on undergraduates, has found that students use notes (Palmatier & Bennett, 1974; Van Meter, Yokoi, & Pressley, 1994), cryptic written records of important information presented in lecture or text, to help them manage their information processing burdens. Palmatier and Bennett (1974), for example, estimated that 99% of undergraduates take lecture notes and 71% take text notes.

Research on the efficacy of notetaking supports the efforts students expend in taking notes. Students who take and/or review lecture notes (Armbruster, 2009; Fisher & Harris, 1973; Kiewra & Benton, 1988; Kiewra, Benton, & Lewis, 1987; Kiewra et al., 1991; Peverly et al., 2007) or text notes (Bretzing & Kulhavy, 1981; Peverly, Brobst, Graham, & Shaw, 2003; Rickards & Friedman, 1978; Slotte & Lonka, 1999) do better on tests than those who do not.

Despite the prevalence and importance of notetaking, there has been very little research on the cognitive variables that underlie skill in any form of notetaking. Cognitive task analyses of *lecture* notetaking (Kobayashi, 2005; Peverly et al., 2007; Piolat et al., 2005) and *text* notetaking (Piolat et al., 2005) suggest that both are cognitively demanding skills. In both, students must comprehend verbal material, hold information in verbal working memory, decide which information is the most important to retain, integrate the information with relevant knowledge (Peper & Mayer, 1986), transcribe it, via writing or typing, before the information is forgotten, and continue to attend to the lecture or text for long periods of time. Thus, skill in notetaking may be related to five variables: comprehension ability (reading or listening), verbal working memory, background knowledge, transcription fluency, and attention.

One might logically infer that some variables are more important to one form of notetaking than the other. For example, because of its limited temporal nature, transcription fluency and working memory capacity might be more important to lecture than text notetaking. However, differences between types of notes have not been addressed explicitly in cognitive task analyses or in previous research (however, see Piolat, 2007).

*Correspondence to: Stephen T. Peverly, Teachers College, Columbia University, Box 120, 525 West 120th Street, New York, NY 10027, USA. E-mail: stp4@columbia.edu

Of the studies that exist on the cognitive underpinnings of notetaking, all but two (Peverly et al., 2003; Piolat, 2007) have focused on lecture notes (e.g. Cohn, Cohn, & Bradley, 1995; Hadwin, Kirby, & Woodhouse, 1999; Kiewra et al., 1987; Kiewra & Benton, 1988; McIntyre, 1992, Peverly et al., 2007). Because of the paucity of research on text notetaking and the lack of theoretical or empirical guides as to which variables might be important to text as opposed to lecture notes, the primary purpose of this investigation was to evaluate the contributions of all the variables identified in the cognitive task analysis discussed above to text notetaking: comprehension, verbal working memory, background knowledge, transcription fluency, and attention. Our ultimate hope, subsequent to many future studies, is to demonstrate that the development of skill of notetaking, lecture, or text, fits the same limited capacity theoretical structure as other basic academic skills. Because cognitive task analyses indicate that notetaking requires the simultaneous execution of at least two skills, theories of limited capacity processing suggest that one of them must be automatized to enable the efficient application of the other skills. As will be discussed later, the ability to take high-quality notes is most likely predicated in transcription speed (Peverly et al., 2007), much as reading comprehension is predicated in the speeded recognition of words (Perfetti, 1985, 2007). This, then, should reduce the burden on working memory, which in turn, should enable the application of higher level cognitive skills like language ability and background knowledge to create high-quality notes.

The second purpose of the experiment was to determine the contribution of all of the aforementioned variables to students' performance on three different tests commonly used in classrooms: two types of multiple choice items, memory and inference, and a written free recall summary (we argue later that a summary is a measure of memory). Although we expected text notes to be significantly related to all three outcome measures, our intent was to determine whether the strength of the relationship would vary with the construct the test was measuring, memory or inference. Prior research has been equivocal. Although some found that *lecture* notes predict performance on measures of memory better than measures of learning (e.g. Kiewra, Dubois, Christensen, Kim, & Lindberg, 1989; Slotte & Lonka, 1999) others found the opposite pattern (Peper & Mayer, 1986; Peverly et al., 2003) or that notes predict both relatively well (Kiewra & Benton, 1988; Kiewra et al., 1987; Williams & Worth, 2003). Yet others have found that performance on memory versus learning measures depends on the structure of notes participants were given to review (e.g. complete text vs. an outline; Kiewra, Dubois, Christian & McShane, 1988), asked to take (e.g. conventional vs. outline; Kiewra, Benton, Kim, Risch, & Christensen, 1995; Kiewra et al., 1988), or the content of notes (factual vs. more conceptual note-takers, Kiewra & Fletcher, 1984). For comprehensive reviews on the relationship of notes to different outcome variables see Armbruster (2009) and Kobayashi (2005). Finally, we wished to determine whether variables other than text notes are related to test performance.

READING COMPREHENSION ABILITY

Reading comprehension refers to readers' mental representation of a text (Kintsch, 1998; Kintsch & Rawson, 2005). To construct a representation, readers process text both locally and globally. At the local or sentence level, readers construct microstructures that are representations of relationships between characters, concepts, actions, goals, etc. Simultaneously, readers construct macrostructures or the gist of the text by using rules that abstract global themes or topics from microstructures and organize them into a hierarchy of superordinate and subordinate relationships.

Logically speaking, individual differences in reading comprehension should be related to individual differences in the amount of important information represented in text notes, with more skilled readers generating higher quality notes than low skilled readers. Unfortunately, there is very little research on the cognitive skills that underlie text notes and none on the relationship of reading comprehension skill to text notes. Three studies on college undergraduates, however, evaluated the relationship of reading comprehension to *lecture* notetaking (Kiewra et al., 1987; Kiewra & Benton, 1988; Peverly et al., 2007). Peverly et al. (2007) failed to find a relationship between *lecture* notes and performance on a text-based main idea identification task constructed by the authors. Also, Kiewra and colleagues (Kiewra et al., 1987; Kiewra & Benton, 1988) failed to find a relationship between lecture notes and performance on the English and Comprehension subtests of the American College Test. Although Peverly et al. (2007) cited methodological flaws with the main idea task as a possible explanation for the lack of a relationship; the American College Test subtests focus directly on the comprehension of written texts. Thus, it is not clear why the latter result was obtained because reading ability and listening comprehension are highly correlated (approximately,9; Gernsbacher, Varner, & Faust, 1990). However, because text notetaking requires students to read and comprehend text, we evaluated the relationship between reading ability and text notes with the Nelson–Denny, a commonly used measure of reading ability for adults.

DOMAIN KNOWLEDGE

Declarative and procedural knowledge systems (Anderson, 1995) are often used to characterize individual differences in domain knowledge. Declarative knowledge is a hierarchical network of concepts. Individual differences are related to the number of concepts represented, the density and strength of the relationships among concepts, and the extent to which lower level concepts are embedded in higher level ones (Chi, Feltovich, & Glaser, 1981; Kintsch, 1998). Procedural knowledge consists of rules for the application of knowledge (i.e. if/then rules). High knowledge individuals have more sophisticated, efficient rules of application than low knowledge individuals.

Ericsson and Kintsch's (1995; Kintsch, 1998) model of long-term working memory explains how domain knowledge impacts readers' comprehension of text. According to the model, concepts from text that are available in working memory act as retrieval cues for text-related information in long-term memory. Readers with a richer, more fully elaborated knowledge base are better than low-knowledge readers at using text cues to access and apply information from long term memory to interpret and remember the content of text, generate inferences to bridge gaps in the coherence of the text, and most generally, to build a macrostructure or summary of the text.

This investigation evaluated the relationship of background knowledge (and other independent variables) to two dependent variables: notes' quality and performance on tests. Regarding the former, it is not clear if background knowledge is related to text notes, regardless of the relationship of background knowledge to the construction of a macrostructure while reading a text. On the one hand, because notes typically focus on main ideas (Armbruster, 2009), and domain knowledge is positively related to the identification/construction of main ideas, Kintsch's theory would predict that high-knowledge students would be able to take quantitatively and qualitatively better notes from domain related texts and lectures than low knowledge students because of their ability to generate a qualitatively superior macrostructure. On the other hand, an ethnographic study of college students' perceptions of the role and processes of *lecture* notetaking (Van Meter et al.) suggested the opposite. Students stated that they often took fewer notes when they were familiar with the content of the course. Relatedly, Peverly et al. (2003) found that background knowledge was not related to the number of macropropositions that college students included in text notes (macropropositions are the logical or rhetorical relationships among propositions that describe the thematic structure of text), even though they had as much time as they desired to take notes.

In contrast to the relationship of domain knowledge to notes, Kintsch's (1998) theory makes very specific predictions about the relationship between domain knowledge and test performance. Because high-knowledge readers can construct a richer macrostructure than low-knowledge students, the theory predicts that high-domain knowledge students will perform better on domain relevant texts than low-domain knowledge students. The theory's prediction has been consistently upheld (McNamara, 2001; McNamara, Kintsch, Songer, & Kintsch, 1996; Recht & Leslie, 1988; Voss & Silfies, 1996; Walker, 1987).

In addition, Kintsch's theory makes specific predictions about the kinds of test questions that will show the largest effects for domain knowledge. Because domain knowledge helps students generate inferences to bridge the gaps in the coherence of texts and verbal communications (e.g. lectures), Kintsch predicts that high-knowledge students will perform substantially better on inference questions than low-knowledge students. The differences between groups on questions that do not require inferences, such as questions on information stated explicitly in text, will be smaller because students do not necessarily have to understand information to remember it (e.g. statistics formulas). Research has supported these predictions (Adams, Bell, & Perfetti, 1995; McNamara & Kintsch, 1996; Peper & Mayer, 1986, Peverly et al., 2003; Voss & Silfies, 1996).

In summary, it is not clear whether background knowledge will be related to notes. However, evidence indicates that domain knowledge will be positively related to students' performance on tests and those differences between high-knowledge and low-knowledge groups will be more pronounced on inference than memory questions (Kintsch, 1998). To evaluate the latter, we included three outcomes measures commonly used in schools. Two of these assessed memory and one assessed inferences. (See *outcome measures of text comprehension* for more information.)

VERBAL WORKING MEMORY

Working memory refers to a cognitive workspace where information from the environment and long term memory is held, manipulated and interpreted, to achieve goals such as remembering, learning, and/or problem solving (Baddeley, 2001). Research suggests that individual differences in working memory are significantly related to individual differences in reading (Baddeley, 2001; Daneman & Carpenter, 1980, 1983; Just & Carpenter, 1992), writing (Kellogg, 2001; Swanson & Berninger, 1996), and mathematics (Swanson & Kim, 2007).

Text notetaking is very cognitively challenging because of the need to keep large amounts of verbal information active in working memory for the purposes of interpretation and transcription (Piolat et al., 2005). Thus, individual differences in verbal working memory may be related to individual differences in text notetaking. Regardless, research on the relationship of working memory and lecture notetaking (there is no research on the relationship of text notes and working memory) has produced mixed results. Some studies have found a significant positive relationship between working memory and note quality (Kiewra et al, 1987; Kiewra & Benton, 1988; McIntyre, 1992) whereas others have not (Cohn et al., 1995; Peverly et al., 2007).

The lack of consistent outcomes between working memory and notes may be because of differences among studies in the measures used to evaluate it. Kiewra and colleagues (Kiewra & Benton, 1988; Kiewra et al., 1987) and McIntyre (1992) used tasks that required the participants either to unscramble randomly ordered words to make a sentence or to arrange randomly ordered sentences to make a coherent paragraph.[1] The materials were always visually present. Complex span tasks, which are the tasks of choice in working memory research, require the participants to process information while it is visually present, but to remember some of the information after it has been removed from view. Because the tasks used by Kiewra and colleagues (Kiewra & Benton, 1988; Kiewra et al., 1987) and McIntyre (1992) do not require the participants to remember and process information in the same way as complex span tasks, they may not adequately measure either storage or processing as they are typically conceived in the working memory literature. Because the research referred to previously on the significant relationship between working memory and achievement has used complex span tasks, we used Daneman and Carpenter's (1980) reading span task,

[1] Some authors refer to these tasks as 'information processing' tasks (e.g. McIntyre, 1992) and others have referred to them as both working memory and information processing tasks (e.g. Kiewra et al., 1987).

versions of which have been used extensively in working memory research with adults and children.

EXECUTIVE ATTENTION

According to Peverly et al. (2007), a possible reason for the failure of verbal working memory to predict lecture note quality or test performance was that prior research may have measured the wrong aspect of working memory. Engle (2002) proposed that individual differences in working memory are related to differences in executive control of attention, which monitors the allocation of cognitive resources while performing a task. Such control overcomes distractions in order to keep information active in working memory. Engle and colleagues have found that individuals high in working memory capacity show greater (i) resistance to proactive interference, where previously learned information causes forgetting of newly learned information (Kane & Engle, 2000); and (ii) inhibition, or the capacity to resist interference from task irrelevant events, such as the ability to ignore extraneous noise during lecture (Conway, Cowan, & Bunting, 2001; Kane, Bleckley, Conway, & Engle, 2001). One task used extensively to measure the inhibition component of executive attention (Kane & Engle, 2003) is the Stroop (Stroop, 1935), which requires subjects to respond to the color a word is printed in (e.g. red), rather than to the word itself, which is a competing color word (e.g. green).

An adapted version of the Stroop was used in the current study. If executive attention underlies notetaking skill, this may explain the mixed findings of previous research on the relationship between working memory and note quality, as executive attention may have had varying roles in the working memory tasks used in those studies.

TRANSCRIPTION FLUENCY

McCutchen (1996) suggested that some of the differences in the quality of children's writing are attributable to insufficient transcription fluency or the ability to write quickly, which, within a limited capacity system, poses difficulty for the efficient and effective use of higher-level processes (e.g. planning). A growing body of research suggests a positive and significant relationship between transcription fluency and different types of writing outcomes for individuals of all ages. Whether focusing on individual differences in transcription fluency (all ages), instruction in transcription fluency (children only) or the experimental manipulation of handwriting speed (via the manipulation of writing scripts; adults only), research with elementary and middle school children (Berninger et al., 1997; Christensen, 2004; Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Graham, Harris, & Fink, 2000; Jones & Christensen, 1999; Olinghouse & Graham, 2009; Olive, Favart, Beauvais, & Beauvais, 2008) and adults (e.g. J. S. Brown, McDonald, Brown, & Carr, 1988; Connelly, Campbell, MacLean, & Barnes, 2006; Connelly, Dockrell, & Barnett, 2005; Olive, Alves, & Castro, 2009; Olive & Kellogg, 2002) has found a significant relationship between transcription fluency and ratings of the quantity and quality of the participants' essays.

Peverly et al. (2007) extended the importance of transcription fluency to lecture notes. They investigated the relationship of several variables to quality of notes (main idea identification, verbal working memory, spelling skill, speed of semantic access, and transcription fluency, among others) and the relationship of all of these to students' performance on a written summary. In two experiments, notes' quality was the only significant predictor of test performance and transcription fluency was the only significant predictor of notes (encoding). We should note that Peverly et al., Connelly and colleagues, and many of the researchers who have focused on children have used the same task to measure transcription fluency; the participants were required to write the letters of the alphabet as many times as they could within a short period of time (e.g. a minute).

Although one could argue that transcription fluency is less important to text notes than to lecture notes (Kobayashi, 2006), both lecture and text notetaking are effortful activities (Piolat et al., 2005) and previous research with children and adults has shown that transcription fluency is important to writing essays and taking lecture notes. Thus, we evaluated the contribution of transcription fluency to the quality of text notes.

OUTCOME MEASURES OF TEXT COMPREHENSION

According to Kintsch (1998), *memory* for text is the ability to recall or recognize information stated directly in the text. Memory does not, however, imply understanding. For instance, taken to the extreme, a person can recall a text verbatim without any understanding of what the text means. Based on this conceptualization, memory in the current study was assessed via written free recall and multiple-choice questions derived from information stated explicitly in text (memory multiple choice, hereafter referred to as Memory MC). In contrast, *inferences* which measure the ability to go beyond information stated explicitly in text imply understanding of the text (Kintsch, 1998). Thus, we included inference-based multiple choice questions (inference multiple choice, hereafter referred to as Inference MC) to assess students' understanding.

SUMMARY

The primary purpose of this study was to identify the cognitive variables underlying skill in text notetaking among undergraduates. This study also sought to establish the extent to which all of the cognitive variables included in this study, including notetaking, contributed to predict test outcomes of various types. Results from this study may help to address the critical issue of identifying the skills that underlie a frequently used and effective study skills technique, which in turn, could lead to better designed intervention programs and better recommendations about the usefulness of notes for specific kinds of tests.

METHOD

Participants

The participants were undergraduate students ($N = 138$) in introductory psychology courses at a midsize university in the northeastern United States of America. They participated for course credit. Their mean age was 19.09 years ($SD = .96$), 55% were women, and 87.9% spoke English as their first language. The race/ethnicity of the sample was diverse: White people (63.6%), African American/Black people (11.4%), Asian American/Asian (6.4%), Latino/a (10.7%), and other (7.8%).

Materials

The materials consisted of a fictitious history text, essay and multiple-choice comprehension tests, measures of reading comprehension, general history knowledge, transcription fluency, executive attention, and verbal working memory. All measures were group administered. Inter-rater agreement in scoring was established for all measures by randomly choosing 10 protocols (14%) to be scored by two different raters. Disagreements were settled by consensus.

History text
The text used in this study was developed by Voss and Silfies (1996). It is an account of the mounting tension between two fictional countries, Anchad and Boxgrave. A fictitious text was used to investigate the effect of general history knowledge on notetaking, rather than specific knowledge of a circumscribed event (Brown, 2005; Voss & Silfies, 1996). Voss and Silfies produced two versions of this text. The 'unexpanded' 773-word version contained key events, which were not linked causally. The 'expanded' 1,386-word version contained all elements of the unexpanded text along with the causal links. The expanded version was used in the current study because its style and degree of detail are more typical of college textbooks. The text consists of 12 paragraphs and was assessed to be at the 12th grade reading level (Flesch–Kincaid Grade Level score, Microsoft Word 2003). In order to make the text similar to that of media undergraduates typically study, the format was changed to better resemble a textbook, magazine, or newspaper article. Specifically, the text was presented in three columns with subtitles and four images (a map showing the location of the two countries and three pictures of important politicians). After these adjustments, the readability remained the same but the word count increased to 1,423 words.

Free recall
The participants were provided with one sheet of paper and told—'In the first session you read a text entitled *Precursors of Anchad-Boxgrave Conflict*. You will now have ten minutes to write an organized summary of that text. Your summary should not extend past the front and back of the next sheet of paper. If you need scrap paper, you may use this directions sheet. However, information on this sheet will not be scored. You will not be able to refer back to your notes. Please write as legibly as you can. The experimenter will periodically remind you of the time as you are working.'

Scoring of notes and recall
The scoring procedure used in this study is based on a procedure used by Peverly et al. (2003). A list of all potentially important macropropositions in the text was compiled by the second author, resulting in 45 statements (the macropropositions are defined as the logical or rhetorical relationships among propositions that describe the thematic structure of text). The relative importance of these statements was subsequently evaluated by three persons, two with a background in history. One had a master's degree in history and the second had a bachelor's degree in history and political science. These raters categorized each of the 45 statements as either 'integral,' 'supportive,' or 'trivial' for an accurate understanding of the text. Raters were also encouraged to identify any additional statements that they deemed important but were not included on the original list of 45 statements. Ten statements were rated as trivial by at least two raters and integral by no raters. These 10 items were excluded from the final scoring rubric, which consisted of 35 statements. All of the raters agreed that the remaining 35 statements were integral.

Given variations in the participants' representations of these statements in notes and in the essay, credit was awarded for entries that did not alter the original meaning of the statement. Likewise, spelling was not penalized unless it changed the meaning of the statement. Credit was not awarded for statements that were imbedded in an incorrect context or for statements that were in a context different from that of the text. For example, credit was not given if fact about Anchad was incorrectly associated with Boxgrave in notes or recall. The total scores for both note quality and essay quality could range from 0 to 35. Two raters, in addition to the second author, scored the protocols. The training consisted of (i) reading the stimulus text; (ii) reviewing the 35 statements; (iii) rereading the text; (iv) practicing the scoring system with three protocols; and (iv) discussing the scoring of the practice protocols. Inter-rater agreement in scoring 10 randomly chosen protocols was .98.

Background knowledge test
Like Voss and Silfies (1996) and Brown (2005), a test of general history knowledge was developed to assess the participants' knowledge of historical events and concepts. This instrument measured general knowledge of history rather than circumscribed knowledge of a particular event. Voss and Silfies found a significant positive relationship between performance on a measure of general history knowledge and understanding (based in inferences) of fictitious history texts. Brown did not find the same relationship; however, a floor effect may have precluded such findings.

Pilot testing was conducted with 21 graduate students and similarly-aged participants to assess the adequacy of 29 multiple choice questions. These questions are based on information found in a high school World History textbook (Gordon, 1996) and were chosen to represent a broad array of historical topics and time periods. Each question was followed by four answer choices, with only one correct answer. Pilot testing resulted in the elimination of nine items. Two questions were eliminated based on feedback by the participants who provided plausible alternative answers to the questions. Six other questions were eliminated because they were not challenging enough, with 95% to 100% of the participants responding correctly. The next most frequently correct item (86%) was also eliminated to bring the test to an even 20 questions.

Reading comprehension test

Reading comprehension ability was assessed with the comprehension component of the Nelson–Denny Reading Test, Form G (Brown, Fishco, & Hanna, 1993). The Nelson–Denny is a widely used, easily administered standardized reading test that is useful for screening reading skills, particularly in older students (Murray-Ward, 1998). The comprehension component contains seven reading passages and 38 multiple-choice comprehension questions. Each question is followed by five possible answers. Under standard administration procedures, 20 minutes are allotted for this test. Brown (2005), however, allowed only 15 minutes to complete the test to increase the amount of variance in participants' performance, thereby better discriminating between those with good and poor comprehension (C. Perfetti, personal communication, February 3, 2003). The current study employed the same procedure. The dependent variable of reading comprehension skill is represented by the participants' raw scores on this task, with a possible range of 0–38.

Multiple-choice comprehension test

The multiple-choice comprehension test used in this study was based on the one developed by Brown (2005). Brown's test consisted of 18 questions with four possible answers for each question. Based on Kintsch's (1986) model of text comprehension, Brown developed nine text-explicit items and nine text-implicit items. The text-explicit questions required recall of information stated directly in the text. The text-implicit items required inference generation based on information in the text and/or domain knowledge. In addition to Brown's nine text-explicit items, which were included in the current test, an additional item was added to bring the total to ten. Of Brown's nine text-implicit items, five were included in the current test and five additional items were developed, bringing the total to ten. This change was intended to increase the inference requirement of the questions, which appeared low in four of Brown's original questions. This decision was also based on the unexpected finding by Brown that performance on text-implicit questions was more related to reading comprehension ability than to history knowledge, which is inconsistent with previous research (Voss & Silfies, 1996). All questions were randomly ordered, as were the four possible answers following each question. The possible score range for each type of question was from 0 to 10.

Students were told 'You will now take a multiple-choice test based on the reading of the passage. Please circle the letter next to the answer you choose. Some questions are based on information stated directly in the reading passage, while others require you to make inferences. Please try your best on all items. Please do not look at your notes during this task. The experimenter will periodically remind you of the time while you are working.'

Transcription fluency

The alphabet task, which was used to measure transcription fluency, was based on the one used by Berninger, Mizokawa, and Bragg (1991) that asked children to write as many letters of the alphabet as they could in 30 seconds (hereafter referred to as *transcription fluency*). In this study, the participants were allotted 45 seconds and were instructed to write the alphabet horizontally in capital letters on a blank sheet, starting with *A*. Once finished, they were to rewrite the entire alphabet in lowercase letters, then again in capital letters, and so on, until the time limit expired. The transcription fluency score equaled the total number of recognizable letters.

Executive attention

Although numerous versions of the original Stroop task have been developed (see Lezak, Howieson, & Loring, 2004), the essential task requirements have remained the same. Individuals are asked to rapidly identify the ink color in which incongruous color-names are printed (i.e. say 'blue' when the word 'red' is written in blue ink). Understood as a measure of selective attention and cognitive flexibility, Spreen and Strauss (1998) noted that the Stroop task assesses 'the ease with which a person can shift his or her perceptual set to conform to changing demands and suppress a habitual response in favor of an unusual one' (p. 213).

The Group Stroop is a paper-and-pencil measure adapted from the Stroop Color and Word Test developed by Golden (1978). Golden's Test is individually administered and requires the participants to respond orally. The Group Stroop, like the test developed by Golden, consisted of word reading, color naming, and incongruent color-word naming.

All conditions began with task-specific instructions, sample items, and practice items. The next two pages of each condition consisted of nine rows of five stimulus items each, totaling 45 items per page. Each test item required the participants to examine a stimulus on the left side of a box, make a judgment about the stimulus based on task-specific instructions, and circle one of three possible response choices in the right half of the box (i.e. red, blue, or green). On the word reading trial, the participants were presented with a black, boldface color name on the left and were instructed to 'circle the word on the right that matches the word in bold on the left.' On the color naming trial, the participants were presented with a red, blue, or green boldface 'XXX' on the left and were instructed to 'circle the name of the color on the right that matches the color on the left.' On the incongruent color-word trial, which is typically considered to be the measure of executive attention in the Stroop, the participants were presented with a boldface color name printed in red, blue, or green ink and were again instructed to 'circle the name of the color on the right that matches the color of the ink on the left.' The participants were given 45 seconds to complete each condition. They were asked to work as quickly as possible until time expired. The score for each of the three trials was the total number of items answered correctly, with a range from 0 to 90.

To evaluate the validity of the Group Stroop, 13 graduate students were administered both the Group Stroop and the Golden's Stroop Color and Word Test (1978). The Group

Stroop was administered first, followed by the individual administration of the Golden Stroop at least 10 days later. The validity of the Group Stroop was assessed by correlating the scores for the first three conditions with scores on the three conditions of Golden's Stroop. Pearson correlations revealed significant positive correlations between the word reading conditions ($r = .61$, $p = .027$), but not the color naming conditions ($r = .33$, $p = .267$). Most importantly, however, there was a strong positive correlation between the incongruent color-word conditions ($r = .89$, $p = .000$). Executive attention was operationalized as the percent decline from color naming to incongruent color-word naming (incongruent color-word/color naming), which is consistent with estimates of interference described elsewhere (Spreen & Strauss, 1998).

Verbal working memory

Verbal working memory was assessed with an adapted version of the Reading Span Task (Daneman & Carpenter, 1980). This version consisted of 40 sentences grouped into two sets of sentences for each of five difficulty levels. Sentences were presented via PowerPoint projected onto a screen using an LCD projector. The first difficulty level consisted of two sets of two sentences each. The next consisted of two sets of three sentences, and so on until the last set, which consisted of two sets of six sentences each. As the participants read each sentence, they were to indicate whether the sentence made sense by circling 'Yes' or 'No' in their response packets. They were also required to remember the last word of each sentence. The end of each set was signaled by a tone and a slide instructing the participants to 'Write Words.' Depending on the number of sentences in the set, the participants had 10–18 seconds to recall and write down the last word from each sentence in the set. When time was up, the participants heard another tone, signaling the beginning of the next set of sentences.

Following the general procedures laid out in Daneman and Carpenter (1980), scores were based on the highest level (2–6) at which the participants remembered all of the words for at least one of the two sets of sentences. More specifically, if a participant correctly recalled all of the final words for both of the sentence sets at level 4, but none at level 5 or 6, his/her score was 4. If a participant correctly recalled all of the words for only one set at level 4, the score was the number of sentences in that set minus .5 (i.e. 3.5). The range of possible scores was from 1.5 to 6, with increments of .5.

The original task developed by Daneman and Carpenter consisted of 60 unrelated sentences grouped into three sets of sentences for each of the five difficulty levels. The primary purpose of modifying their task was to reduce its length in an effort to lessen the participants' fatigue. Pilot testing with 17 graduate students revealed a correlation of .98 between the original and adapted procedures.

Procedure

The participants completed all tasks in a group format over the course of two 55-minute sessions held on two separate days. The intersession interval was between seven and nine days. All the participants provided informed consent to participate at the start of the first session. The participants were then given 15 minutes to take notes on the history text. They were informed that they would be studying these notes during the second session to prepare for an essay test and a multiple choice comprehension test. They were also informed that, because they would have only their notes to study from, it was important that they take the best notes possible. The order of the remaining procedures for the first session was: Group Stroop, Transcription fluency, and Nelson–Denny Reading Test. At the start of the second session, the participants completed the general history knowledge test, followed by 10 minutes to study their notes from the previous session. They were reminded of the essay test and multiple choice comprehension tests that would be administered later in the session. Following this study period, the participants completed the reading span task. Next, they were asked to compose an essay about the events leading up to war between Anchad and Boxgrave. Finally, the participants completed the multiple-choice comprehension test. They were informed that half of the items required knowledge of information stated directly in the text, whereas the other half required them to generate inferences based on information in the text and/or their own knowledge.

RESULTS

Means and standard deviations for all variables are in Table 1, and intercorrelations among variables are in Table 2.

First, to evaluate the primary purpose of the study, note quality was regressed on each of the independent variables. The regression equation was significant (tolerance and variance inflation factor values were within acceptable limits; $R = .58$, $R^2 = .34$, $R^2_{\text{adjusted}} = .31$), $F(5, 137) = 13.46$, $p < .001$ (the effect size was large). Transcription fluency ($\beta = .36$, $p < .001$,) was the strongest predictor followed by reading comprehension ability ($\beta = .25$, $p < .01$). See Table 3.

All subsequent analyses evaluated the second purpose of the study, the contribution of all of the cognitive variables, including notes, to test outcomes. For essay quality, the equation was significant (tolerance and variance inflation factor values were within acceptable limits; $R = .70$, $R^2 = .49$,

Table 1. Means and standard deviations

Statistic	Note quality	Essay quality	Memory MC	Inference MC	Read comp	Hist know	Trans flu	Exec attn	VWMC
M	18.15	11.01	6.45	5.26	25.36	9.99	82.64	0.76	3.75
SD	5.31	5.44	1.98	1.77	5.57	3.26	14.77	0.11	1.52

Note: MC, multiple choice; Read comp, reading comprehension; Hist know, history knowledge; Trans flu, transcription fluency; Exec attn, executive attention; VWMC, verbal working memory capacity.

$R^2_{adjusted} = .46$), $F(6, 137) = 20.57$, $p < .001$ (the effect size, with R^2 used as an estimate of effect size, was large; Cohen, 1992). Notes ($\beta = .58$, $p < .001$) were the only significant predictor of essay quality. See Table 4.

Next, Memory MC performance was regressed on all of the independent variables. The equation was significant (tolerance and variance inflation factor values were within acceptable limits; $R = .61$, $R^2 = .37$, $R^2_{adjusted} = .34$), $F(6, 137) = 12.75$, $p < .001$ (the effect size was large). The best predictor of Memory MC performance was notes ($\beta = .32$, $p < .001$) followed by history knowledge ($\beta = .25$, $p < .01$) and reading comprehension ability ($\beta = .17$, $p < .05$). See Table 5.

In the third analysis, Inference MC performance was regressed on all the independent variables. The equation was significant (tolerance and variance inflation factor values were within acceptable limits; $R = .48$, $R^2 = .23$, $R^2_{adjusted} =$

Table 2. Intercorrelations among the independent and dependent variables

Variables	1	2	3	4	5	6	7	8	9
1. Note quality	—								
2. Essay quality	.66***	—							
3. Memory MC	.50***	.56***	—						
4. Inference MC	.27***	.44***	.38***	—					
5. Read comp	.41***	.41***	.43***	.35***	—				
6. Hist know	.28***	.33***	.41***	.40***	.45***	—			
7. Trans flu	.45***	.28**	.26**	.04	.20*	.04	—		
8. Exec attn	.18*	.05	.16	.15	.09	−.08	.21*	—	
9. VWMC	.22**	.28***	.13	.06	.18*	.15	.19*	.05	—

Note: MC, multiple choice; Read comp, reading comprehension; Hist know, history knowledge; Trans flu, transcription fluency; Exec attn, executive attention; VWMC, verbal working memory capacity.
*$p < .05$, **$p < .01$, ***$p < .001$.

Table 3. Summary of the regression analysis predicting note quality

Variable	*B*	*SE B*	β	Partial *r*	Tolerance	VIF
Exec attention	4.42	3.42	.09	.11	.94	1.06
Trans flu	0.13	0.03	.36	.38***	.90	1.11
Reading comp	0.24	0.08	.25	.26**	.75	1.34
History know	0.24	0.13	.15	.16	.78	1.29
VWM	0.29	0.26	.08	.10	.90	1.07

Note: $R = .58$, $R^2 = .34$, $R^2_{adjusted} = .31$. VIF, variance inflation factor; Trans flu, transcription fluency; Exec, executive; comp, comprehension; know, knowledge; VWM, verbal working memory; VIF, variance inflation factor; *SE*, standard error.
$p < .05$, **$p < .01$, ***$p < .001$.

Table 4. Summary of the regression analysis predicting essay quality

Variable	*B*	*SE B*	β	Partial *r*	Tolerance	VIF
Note quality	0.59	0.08	.58	.55***	.66	1.51
Exec attention	−2.72	3.13	−.06	−.08	.93	1.08
Trans flu	−0.01	0.03	−.02	−.03	.77	1.30
Reading comp	0.11	0.07	.11	.13	.70	1.43
History know	0.16	0.12	.09	.11	.76	1.32
VWMC	0.45	0.23	.12	.17	.93	1.08

Note: $R = .70$, $R^2 = .49$, $R^2_{adjusted} = .46$. VIF, variance inflation factor; Exec, executive; Trans Flu, transcription fluency; comp, comprehension; know, knowledge; VWMC, verbal working memory capacity; *SE*, standard error.
***$p < .001$.

Table 5. Summary of the regression analysis predicting memory multiple choice performance

Variable	*B*	*SE B*	β	Partial *r*	Tolerance	VIF
Note quality	0.12	0.03	.32	.31***	.66	1.51
Exec attention	1.75	1.26	.10	.12	.93	1.08
Trans flu	0.01	0.01	.06	.06	.77	1.30
Reading comp	0.06	0.03	.17	.18*	.70	1.43
History know	0.16	0.05	.26	.27**	.76	1.32
VWMC	−0.04	0.09	−.03	−.03	.93	1.08

Note: $R = .61$, $R^2 = .37$, $R^2_{adjusted} = .34$. VIF, variance inflation factor; Exec, executive; Trans flu, transcription fluency; comp, comprehension; know, knowledge; VWMC, verbal working memory capacity; *SE*, standard error.
*$p < .05$, **$p < .01$, ***$p < .001$.

.20), $F(8, 137) = 6.60$, $p < .001$ (the effect size was large). The only significant predictor of Inference MC performance was history knowledge ($\beta = .30$, $p < .001$). See Table 6.

DISCUSSION

The primary purpose of this study was to identify some of the cognitive variables underlying text notetaking skill among undergraduates. A secondary purpose was to establish the importance of text notes as a predictor of performance on the outcome measures of memory and inferencing, and to identify variables that contribute to performance on these measures over and above notes.

Prediction of text notetaking skill

Consistent with lecture notetaking research (Peverly et al., 2007), the best predictor of text note quality was transcription fluency. Reading comprehension ability was also a strong predictor. None of the other variables were significant.

Consistent with research on written composition (Brown et al., 1988; Olive & Kellogg, 2002; Peverly, 2006), greater transcription fluency likely resulted in greater automaticity of word production, which we hypothesize, lessened the burden on students' working memory and enabled them to use their limited capacity working memory resources more efficiently to produce greater quality text notes. The question is, however, what does transcription fluency, as it was operationalized in this study, actually measure? Given the requirements of both manual movements and speed, it would be natural to assume that fine motor speed underlies transcription fluency. Research suggests, however, that transcription fluency is more strongly related with orthographic coding than fine motor speed (Abbott & Berninger, 1993; Berninger et al., 2006; Berninger & Richards, 2002). In a cross-sectional study of writing skill development in children across first to sixth grades, Abbott and Berninger (1993) found that fine motor speed did not predict transcription fluency directly, but was mediated by orthographic coding. Likewise, Berninger et al. (2006) found that orthographic coding was related to cursive writing in third and fifth grades.

Orthographic coding is a component of a larger verbal system (James & Gauthier, 2006), thereby explaining Peverly et al.'s (2007) finding that transcription fluency (which they labeled letter fluency) was correlated with almost all of the verbal measures used across two experiments, including spelling ability (i.e. orthographic skill). Because of the significant correlations between transcription fluency and verbal measures, Peverly et al. proposed that these tasks may share a common underlying construct, rapid access to phonological codes. This would also explain their significant correlations of .30 and .51 between transcription fluency and two verbal fluency tasks (semantic and phonetic, respectively), which are hypothesized to measure speed of access of information from long-term memory.

The second strongest predictor of note quality was reading comprehension ability. This finding supports discussion of the cognitive variables that underlie skill in lecture notetaking (Peverly et al., 2007; Piolat et al., 2005) that suggest that verbal ability underlies students' ability to detect important information in lecture. Because reading comprehension skill correlates very highly with listening comprehension skill (from .8 to .9), this suggests that verbal ability may underlie the ability to detect important information in both skills (Gernsbacher et al., 1990). Future research should validate this outcome with lecture notes.

Neither executive attention nor verbal working memory capacity contributed to text note quality. Regarding executive attention, Engle (2002) argued that working memory might be related more to executive attention than to the variables measured by a complex span task (capacity + processing). Thus, given that research has not found a significant relationship between verbal working memory and notetaking when using a complex span task of the type we used here, we used the measure of executive attention (the Stroop) used by Engle in his research. However, executive attention is a relatively broad construct and some have argued that the Stroop measures only one aspect of executive attention, response control, and inhibition (Spreen & Strauss, 1998). It may be that response control and inhibition are less important to notetaking than another component of executive attention, sustained attention, which is conceptualized as the ability to maintain attention across a longer period of time (Lezak et al., 2004). In a preliminary analysis of data in one of our current projects on lecture notetaking, transcription fluency and sustained attention were the only significant predictors of the quality of lecture notes. Other variables, such as reading comprehension, verbal working memory, and the Stroop, among other variables, were not significant predictors.

Table 6. Summary of the regression analysis predicting inference multiple choice

Variable	*B*	*SE B*	β	Partial *r*	Tolerance	VIF
Note quality	0.05	0.03	.14	.13	.66	1.51
Exec attention	2.38	1.24	.15	.17	.93	1.08
Trans flu	−0.01	0.01	−.09	−.09	.77	1.30
Reading comp	0.05	0.03	.17	.16	.70	1.43
History know	0.16	0.05	.30	.29***	.76	1.32
VWM capacity	−0.05	0.09	−.04	−.04	.93	1.08

Note: $R = .48$, $R^2 = .23$, $R^2_{adjusted} = .20$. VIF, variance inflation factor; Exec, executive; Trans flu, transcription fluency; comp, comprehension; know, knowledge; VWM, verbal working memory; *SE*, standard error.
***$p < .001$.

The failure of verbal working memory capacity to uniquely predict text note quality is consistent with the research on lecture notetaking that has used verbal working memory tasks (complex span) of the kind used in this study (Cohn et al., 1995; Peverly et al., 2007). As noted above, this may be partially explained by the existence of a common underlying construct between this task and the transcription fluency task, namely, the speeded access to phonological codes. Another consideration is the methods we used to assess executive attention and verbal working memory. More specifically, although both of these tasks were based on measures developed for individual administration, they were adapted for group administration in this study. Although pilot testing found a high correlation ($r = .89$) between the individually and group administered versions of the Stroop, group administration nonetheless may have inadvertently altered the constructs being measured. The Stroop task traditionally requires a vocal response that may elicit different cognitive processes than the nonvocal response in this study. In a group administration of the reading span task, it can be more difficult to monitor the participants' compliance with the requirements of the task (e.g. our instruction not to write down the last word of the last sentence first) or for individual differences in completing the processing component (verification of the meaningfulness of the sentences), which can have an effect on rehearsal times (Conway et al., 2005). In contrast, Cohn et al. (1995) used an individually administered complex span task and did not find a significant relationship between working memory and lecture notes. Also, we just completed a study on lecture notetaking where we administered the listening span task individually and did not find a significant relationship between verbal working memory and notes or performance. These findings suggest that it may not matter whether complex span tasks are presented individually or in groups when evaluating their relationship to notes.

Prediction of test performance

Notes were the most important contributor to undergraduates' memory for text. They were the only significant predictor of students' essay performance and they were the strongest of three significant predictors of Memory MC. In contrast, notes did not significantly predict inferences. These findings are consistent with Kiewra et al. (1989) and Slotte and Lonka (1999) but not Kiewra and Benton (1988), Kiewra et al. (1987), Peper and Mayer (1986) and Peverly et al. (2003), among others. Reasons for the variation in outcomes may be related to variations in study design and materials, the procedures used to score notes, and the definition and operationalization of dependent variables. Relative to the issue of study design, for example, Peper and Mayer (1986) found that there were no significant differences between a notes and no-notes group on memory items but that the notes group was superior to the no-notes group on inference items. However, the design of their study is not directly comparable with ours because they did not analyze notes, we did not include a no-notes group, and they did not evaluate within-group differences on inference and memory items.

Another example of study design and materials differences is Peverly et al. (2003). One of the differences between Peverly et al. (2003) and the current study is the degree to which readers had to use inferencing skills to identify or construct the text's macropropositions. In Peverly et al., all of the causal macropropositions had to be inferred. None were stated directly in text. In the current study, many of macropropositions were stated directly in text and did not require readers to infer relationships. Thus, the macropropositions identified by Peverly et al. in both notes and recall were primarily the measure of students' skill in generating inferences. Because, as was discussed previously, the information included in notes is more likely to be included in recall than information that was not included in notes (Bretzing & Kulhavy, 1981; Kiewra et al., 1987) and the scoring of notes and recall focused on inferences, it makes sense that an inference-based measure of notes' quality was a good predictor of the inferences included in recall. Thus, how notes are operationalized may have an impact on prediction patterns. When notes are operationalized as the representation of information stated explicitly in text, they seem to predict students' performance on measures of memory. When operationalized as inferences, notes seem to predict performance on tests that measure skill in inferencing (e.g. Peverly et al., 2003).

In addition, a variety of procedures have been used to analyze notes over the years resulting in a number of different measures of notes quantity and/or quality: word counts, estimates of notes' local and global coherence, and estimates of the number of idea units, propositions, macropropositions, and/or main ideas, among others. Also, the methods used to determine what a main idea or idea unit are can vary from one study to the next. Thus, part of the inconsistency in the relationship of notes to different outcome measures may be because of the variety of procedures used to analyze notes.

The prediction patterns for each dependent variable are discussed in more detail in the following sections.

Free recall

A great deal of evidence indicates that readers automatically form a macrostructure or global, hierarchical representation of the most important information in text (Kintsch, 1998), and there is some evidence that this tendency is intensified if readers are asked to create an outline of the text (Lorch, Lorch, & Matthews, 1985), which is similar to taking notes. Thus, notes may have held a monopoly on the unique variance in essay quality because written recall may have simply given the students the opportunity to access and write down the macrostructure of the text they encoded during notetaking. Support for this conjecture comes from notetaking research that indicates that students are more likely to recall information recorded in their notes than information that was not recorded (Bretzing & Kulhavy, 1981; Kiewra et al., 1987). Because reading comprehension skill was significantly related to the quality of notes, these findings also suggest that better readers may construct a qualitatively better representation of the text in long term memory (and notes) and have better retrieval cues for that information than do poorer readers (Ericsson & Kintsch, 1995; Kintsch, 1998).

Memory multiple choice

History knowledge and reading comprehension, in addition to notes, predicted performance on Memory MC. Although the correlation (.58) between essay quality and memory MC suggests that both measure some aspects of memory, it also suggests that they are more different than alike ($r^2 = .34$).

One possible reason why reading comprehension skill and background knowledge were related to Memory MC and not Recall is that multiple-choice items can be a form of cued recall, and higher levels of reading comprehension skill and background knowledge may enable students to use cues in the questions more effectively to access information encoded in their macrostructures. Also, because the cues in questions may not always match how information was encoded, reading comprehension and background knowledge may enable students to infer a relationship between them and transfer elements of the macrostructure to novel question formats. Higher levels of reading comprehension (Kintsch & Rawson, 2005; Nation, 2005) and background knowledge (Bransford, Brown, & Cocking, 2000; Bransford & Schwartz, 1999; Fuchs et al., 2003; McNamara, 2001; McNamara et al., 1996; Recht & Leslie, 1988; Voss & Silfies, 1996; Walker, 1987) have been consistently linked with inferences and transfer.

Inference multiple choice

The only significant predictor of inferences was domain knowledge. This finding is consistent with theory (Kintsch, 1998) and previous research in both text comprehension (Kintsch, 1998; McNamara & Kintsch, 1996; Kintsch & Rawson, 2005; Voss & Silfies, 1996) and text notetaking (Peverly et al., 2003), at least in situations where notes and recall are evaluated for students' ability to generate inferences.

Summary

Notes are the strongest predictor of students' performance on measures of memory but domain knowledge and reading comprehension contribute to students' memory for text in situations where the test does not allow students' to simply recall the information they recorded and reviewed in their notes. In other words, in comparison with written recall, Memory MC may be a measure of students' ability to detect cues and transfer their knowledge to novel question formats. Notes do not appear, however, to predict students' performance on inference questions. Consistent with the prior research in text processing, background knowledge was related to inferences.

SUMMARY AND IMPLICATIONS

Very little research has focused on the cognitive skills that underlie effective note-taking (Armbruster, 2009). Of the few studies that do exist on the cognitive processes that underlie notes, as stated previously, all but two (Peverly et al., 2003; Piolat, 2007) focused on lecture notes (e.g. Cohn, Cohn, & Bradley, 1995; Hadwin, Kirby, & Woodhouse, 1999; Kiewra et al., 1987; Kiewra & Benton, 1988; McIntyre, 1992, Peverly et al., 2007). This study added to past research by expanding our understanding of the cognitive processes that underlie text notetaking, by demonstrating that the cognitive processes that underlie text notes may be similar to those that underlie lecture notes and by demonstrating that notes may be more strongly related to some types of tests or test items than others. We explore these issues more in depth in the paragraphs below.

Skill in taking text notes was predicted by transcription fluency and reading comprehension skill (verbal ability). The finding on transcription fluency suggests that it may be a cognitive constant in both text and lecture notetaking. Although more research is needed to support this claim, Peverly et al. (2007) found that transcription fluency was the only variable to predict skill in lecture notetaking (transcription fluency was called letter fluency in their study), and a preliminary analysis of data from two studies on lecture notetaking just completed in our laboratory found that transcription fluency was a significant predictor of notes (along with other variables) in both studies.

If our finding on the relationship of verbal ability to notes is verified in future research on text notes and is extended to lecture notes, the overall pattern of cognitive processing in notetaking could parallel that in other academic skill domains. More specifically, contemporary views of cognitive processing (e.g. Anderson, 1995; Baddeley, 2001; Ericsson & Kintsch, 1995; Schneider & Shiffrin, 1977) suggest that learning and successfully executing many school-based tasks depends on performing a hierarchy of skills simultaneously, in parallel. In the execution of these skills, domain-specific basic skills must be executed automatically (e.g. work recognition) so that most if not all of the space in working memory can be used for the application of the higher level cognitive skills (e.g. language) needed to support comprehension, reasoning, remembering, learning, and the like. If these skills are not automatized, the application of higher level cognitive skills can be attenuated and prevent students from realizing the goal of the activity (e.g. effective comprehension), even if students' cognitive resources are substantial. Translated into the current context, data from this study suggest that transcription must be sufficiently fluent to enable the application of higher order reading/language skills to create high quality text (lecture) notes. Thus, the current study in combination with some prior research may provide the beginnings of a theory of notetaking that parallels those in other academic domains.

We also evaluated the relationship of notes and other independent variables to different test and item types typically used in schools: Essay, Memory MC, and Inference MC. Although research has consistently found that notes are a strong and consistent predictor of test outcomes, the research on whether notes predicts different types of questions/tests equally well is equivocal. Also, there is almost no research on the question of whether variables other than notes combine with notes to predict certain test outcomes.

The results of this investigation support the findings of Kiewra et al. and Slott and Lonka to the extent that notes predict memory performance better than students' performance on inference questions. The current study also extends their research by demonstrating that notes predict performance on some kinds of memory measures better than others. Notes were the only predictor of Recall. However, notes, reading comprehension skill, and background knowledge predicted Memory MC. Although Kintsch (1998) would label Recall and Memory MC as memory tests, because they do not directly test students' understanding via the processes of inference and problem solving, the different prediction patterns may be a result of a mismatch between how information was encoded and tested. Recall is straightforward; students need only write what they encoded. Thus, free recall is probably a good measure of the macrostructure students constructed during notetaking. In contrast, Memory MC may have tested students' ability to transfer what they learned to novel question formats. Thus, reading comprehension skill and background knowledge may be necessary for transfer. Finally, only background knowledge was related to the ability to generate inferences. This finding verifies the relationship between knowledge and inferences found in research (McNamara, 2001; McNamara et al., 1996; Peverly et al., 2003; Recht & Leslie, 1988; Voss & Silfies, 1996; Walker, 1987) and theory (Kintsch, 1998).

Finally, fluency as we have used the word in the context of written transcription implies a linear relationship to performance. That is, more is better. Although that may be true of written transcription, given the general limitations in handwriting speed, it may not be true of other forms of transcription, such as keyboarding. Some students who type very quickly reported to us that they switched from laptops to paper and pencil to take notes. They stated that they understood more of the lecture if they recorded information more slowly, which suggests that the relationship of keyboarding to notetaking may be curvilinear. Research will need to be conducted to evaluate this relationship.

LIMITATIONS

This study has some limitations. First, text notetaking was measured under somewhat unrealistic conditions. Students were given a 15 minute time limit and text notetaking is not usually time limited. This study needs to be rerun without a notetaking time limit to see if the results can be replicated.

Second, we used a group version of the Stroop and the working memory task. Although Cohn et al. (1995) used a complex span task and found that working memory was not related to notes (lecture notes in their case) and another as yet unpublished *lecture* notetaking study from our laboratory used individually administered versions of the Stroop and working memory and found the same nonsignificant relationships between those measures and notes, a text notetaking study should be run using individual administrations of both measures.

REFERENCES

Abbott, R., & Berninger, V. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary and intermediate grade writers. *Journal of Education & Psychology*, *85*, 478–508.

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

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th edn), New York: W. H. Freeman and Co.

Armbruster, B. (2009). Notetaking from lectures. In R. F. Flippo, & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (2nd edn, pp. 220–248). New York: Routledge.

Baddeley, A. D. (2001). Is working memory still working? *The American Psychologist*, *56*, 849–864.

Berninger, V. W., Abbott, R. D., Jones, J., Wolf, B. J., Gould, L., & Anderson-Youngstom, M. (2006). Early development of language by hand: Composing, reading, listening, and speaking-connections; three letter writing modes; and fast mapping in spelling. *Developmental Neuropsychology*, *29*, 61–92.

Berninger, V. W., Mizokawa, D., & Bragg, R. (1991). Theory based diagnosis of remediation of writing disabilities. *Journal of School Psychology*, *29*, 57–79.

Berninger, V. W., & Richards, T. L. (2002). *Brain literacy for educators and psychologists*. San Diego, CA: Academic Press.

Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Rogan, L. W., & Brooks, A. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Education & Psychology*, *89*, 652–666.

Bransford, J., & Schwartz, D. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Jejad, & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 61–100). Itasca, IL: Peacock.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

Bretzing, B. H., & Kulhavy, R. W. (1981). Note-taking and passage style. *Journal of Education & Psychology*, *7*, 242–250.

Brown, C. (2005). An investigation of the influence of metacognition, reading comprehension skill, and background knowledge on studying. (Doctoral dissertation, Columbia University, United States, 2005).

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny reading test*. Chicago, IL: Riverside Publishing Co.

Brown, J. S., McDonald, J. L., Brown, T. L., & Carr, T. H. (1988). Adapting to processing demands in discourse production: The case of handwriting. *Journal of Experimental Psychology. Human Perception and Performance*, *14*, 45–59.

Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.

Christensen, C. A. (2004). Relationship between orthographic-motor integration and computer use of the production of creative and well-structured written text. *The British Journal of Educational Psychology*, *74*, 551–564.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

Cohn, E., Cohn, S., & Bradley, J. (1995). Notetaking, working memory, and learning in principles of economics. *Research in Economic Education*, *26*, 291–307.

Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2006). Contribution of lower order skills to the written composition of college students with and without dyslexia. *Developmental Neuropsychology*, *29*, 175–196.

Connelly, V., Dockrell, J. E., & Barnett, J. (2005). The slow handwriting of undergraduate students constrains overall performance in exam essays. *Educational Psychology*, *25*, 97–105.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The important of WM capacity. *Psychonomic Bulletin & Review*, *8*, 331–335.

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

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

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *9*, 561–583.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245.

Fisher, J. L., & Harris, M. B. (1973). Effect of note taking and review on recall. *Journal of Education & Psychology*, *65*, 321–325.

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., … & Jancek, D. (2003). Explicitly teaching for transfer: Effects on third-grade students' mathematical problems solving. *Journal of Education & Psychology*, *95*, 293–305.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *16*, 430–445.

Golden, J. C. (1978). *Stroop color and word test*. Chicago, IL: Stoelting Co.

Gordon, I. L. (1996). *World history*: 2nd edn, New York, NY: Amsco School Publications, Inc.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Education & Psychology*, *89*, 170–182.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment in handwriting problems in beginning writers. *Journal of Education & Psychology*, *92*, 620–633.

Hadwin, A. F., Kirby, J. R., & Woodhouse, R. A. (1999). Individual differences in notetaking, summarization, and learning from lectures. *The Alberta Journal of Educational Research*, *45*, 1–17.

James, K. H. & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*, *44*, 2937–2949.

Jones, D., & Christensen, C. A. (1999). Relationship between automaticity in handwriting and students' ability to generate text. *Journal of Education & Psychology*, *91*, 44–49.

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

Kane, M. J., & Engle, R. W. (2000). WM capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*, 336–358.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47–70.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled attention view of WM capacity. *Journal of Experimental Psychology: General*, *130*, 169–183.

Kellogg, R. T. (2001). Competition for working memory among writing processes. *The American Journal of Psychology*, *114*, 175–191.

Kiewra, K. A., Benton, S. L. (1988). The relationship between information processing ability and notetaking. *Contemporary Educational Psychology*, *13*, 33–44.

Kiewra, K. A., Benton, S. L., Kim, S.-I., Risch, N., & Christensen, M. (1995). Effects of note-taking format and study technique on recall and relationship performance. *Contemporary Educational Psychology*, *20*, 172–187.

Kiewra, K. A., Benton, S. L., & Lewis, L. B. (1987). Qualitative aspects of notetaking and their relationship with information-processing ability and academic achievement. *Journal of Instructional Psychology*, *14*, 110–117.

Kiewra, K. A., Dubois, N. F., Christensen, M., Kim, S. I., & Lindberg, N. (1989). A more equitable account of the note-taking functions in learning from lecture and from text. *Instructional Science*, *18*, 217–232.

Kiewra, K. A., DuBois, N. F., Christian, D., & McShane, A. (1988). Providing study notes: Comparison of three types of notes for review. *Journal of Educational Psychology*, *80*, 595–597.

Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology*, *83*, 240–245.

Kiewra, K. A., & Fletcher, H. J. (1984). The relationship between levels of note-taking and achievement. *Human Learning*, *3*, 273–280.

Kintsch, W. (1986). Learning from text. *Cognition and Instruction*, *3*, 87–108.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading* (pp. 209–226). Malden, MA: Blackwell Publishing.

Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology*, *30*, 242–262.

Kobayashi, K. (2006). Combined effects of notetaking/-reviewing on learning and the enhancement through interventions: A meta-analytic review. *Contemporary Educational Psychology*, *26*, 459–477.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th edn), New York, NY: Oxford University Press.

Lorch, R. F., Lorch, E. P., & Matthews, P. D. (1985). On-line processing of the topic structure of a text. *Journal of Memory and Language*, *24*, 350–362.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, *8*, 299–325.

McIntyre, S. (1992). Lecture notetaking, information processing, and academic achievement. *Journal of College Reading and Learning*, *25*, 7–17.

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

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

McNamara, D. S., Kintsch, E., Songer, N. B., & 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.

Mulcahy-Ernt, P. I., & Caverly, D. C. (2009). Strategic study-reading. In R. F. Flippo, & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (2nd edn, pp. 177–198). New York: Routledge.

Murray-Ward, M. (1998). [Review of the measure *Nelson-Denny Reading Test*]. In Impara, J. C., & Plake, B. S. (Eds.), *The thirteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.

Nation, K. (2005). Children's reading comprehension difficulties. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading* (pp. 248–265). Malden, MA: Blackwell Publishing.

Olinghouse, N. G., & Graham, S. (2009). The relationship between discourse knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology*, *101*, 37–50.

Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high- and low-level production processes in written composition. *Memory & Cognition*, *30*, 594–600.

Olive, T., Alves, R. A., & Castro, S. L. (2009). Cognitive processes in writing during pauses and execution periods. *European Journal of Cognitive Psychology*, *21*, 758–785.

Olive, T., Favart, M., Beauvais, C., & Beauvais, L. (2008). Children's cognitive effort and fluency in writing: Effects of genre and of handwriting automatisation. *Learning and Instruction*, *19*, 299–308.

Palmatier, R. A., & Bennett, J. M. (1974). Notetaking habits of college students. *Journal of Reading*, *18*, 215–218.

Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology*, *78*, 34–38.

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

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.

Peverly, S. T. (2006). The importance of handwriting speed in adult writing. *Developmental Neuropsychology*, *29*, 197–216.

Peverly, S. T., Brobst, K. E., Graham, M., & Shaw, R. (2003). College adults are not good at self-regulation: A study of the relationship of self-regulation, note-taking, and test-taking. *Journal of Educational Psychology*, *95*, 335–346.

Peverly, S. T., Ramaswamy, V., Brown, C., Sumowski, J., Alidoost, M., & Garner, J. (2007). What predicts skill in lecture note taking? *Journal of Educational Psychology*, *99*, 167–180.

Piolat, A. (2007). Effects of note-taking and working memory span on cognitive effort and recall performance. In G. Rijlaarsdam (Series Ed.) and M. Torrance, L. van Waes, & D. Galbraith (Volume Eds.), *Writing and cognition: research and applications* (Studies in Writing, Amsterdam: Elsevier Vol. *20*, pp. 109–124).

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology*, *19*, 291–312.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*, 31–74.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, *80*, 16–20.

Rickards, J. P., & Friedman, F. (1978). The encoding versus the external storage hypothesis in note taking. *Contemporary Educational Psychology*, *3*, 136–143.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66.

Slotte, V., & Lonka, K. (1999). Review and process effects of spontaneous note-taking on text comprehension. *Contemporary Educational Psychology*, *24*, 1–20.

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

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd edn), New York, NY: Oxford University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Swanson, L., & Berninger, V. (1996). Individual differences in children's working memory and writing skills. *Journal of Experimental Psychology*, *63*, 358–385.

Swanson, L., & Kim, K. (2007). Working memory, short-term memory, and naming speed as predictors of children's mathematical performance. *Intelligence*, *35*, 151–168.

Thomas, J. W., Iventosch, L., Rohwer, W. D. (1987). Relationships among student characteristics, study activities, and achievement as a function of course characteristics. *Contemporary Educational Psychology*, *12*, 344–364.

Van Meter, P., Yokoi, L., & Pressley, M. (1994). College students' theory of note-taking derived from their perceptions of note-taking. *Journal of Educational Psychology*, *86*, 323–338.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*, 2–40.

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

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

Williams, R. L., & Worth, S. L. (2003). Thinking skills and work habits: Contributors to course performance. *The Journal of General Education*, *51*, 200–227.