Journal of Educational Psychology
2007, Vol. 99, No. 4, 888–902

0022-0663/07/$12.00 DOI: 10.1037/0022-0663.99.4.888

The Impact of Print Exposure Quality and Inference Construction on Syllogistic Reasoning

Helena P. Osana
Concordia University

Guy L. Lacroix
Carleton University

Bradley J. Tucker, Einat Idan, and Guillaume W. Jabbour
Concordia University

This study extended the work of S. Siddiqui, R. F. West, and K. E. Stanovich (1998), who studied the link between general print exposure and syllogistic reasoning. It was hypothesized that exposure to certain text structures that contain well-delineated logical forms, such as popularized scientific texts, would be a better predictor of deductive reasoning skill than general print exposure, which is not sensitive to the quality of an individual's reading activity. Furthermore, it was predicted that the ability to generate explanatory bridging inferences while reading would also be predictive of syllogistic reasoning. Undergraduate students ($N = 112$) were tested for vocabulary, nonverbal cognitive ability, exposure to general print, exposure to popularized scientific literature, and the ability to comprehend texts distinguished by the number of inferences that must be generated to support comprehension. Hierarchical multiple regression analyses showed that a combined measure of exposure to general and scientific literature was a significant predictor of syllogistic reasoning ability. Additionally, the ability to comprehend high-inference-load texts was related to solving syllogisms that were inconsistent with world knowledge, indicating an overlap in deductive reasoning skill and text comprehension processes.

Keywords: syllogistic reasoning, print exposure, inference construction, text comprehension, science text

A considerable amount of research now documents the cognitive outcomes of experiences with print, such as orthographic processing (Stanovich & West, 1989); general reading ability, including comprehension (Allen, Cipielewski, & Stanovich, 1992; Guthrie, Wigfield, Metsala, & Cox, 1999); verbal fluency and oral language skills (Sénéchal, LeFevre, Thomas, & Daley, 1998; Stanovich & Cunningham, 1992); vocabulary (Cunningham & Stanovich, 1991; Stanovich & Cunningham, 1992); and knowledge of metalinguistic terms (Siddiqui, West, & Stanovich, 1998). The relationship between reading and the development of broader cognitive processes has also received attention from scholars. For instance, Stanovich and his colleagues (see Cunningham & Stanovich, 1991; Stanovich & Cunningham, 1993; Stanovich, West, & Harrison, 1995) found that print exposure contributed unique variance to the development of knowledge structures and information acquisition.

The research on the relationship between print exposure and reasoning skill, such as syllogistic reasoning and argumentation, has not been as successful, however. As part of a larger study, Siddiqui et al. (1998) tested the hypothesis that general print exposure would be positively correlated to syllogistic reasoning. They administered two instruments to 133 undergraduates: the Author Recognition Test (ART; Stanovich & West, 1989) and the Magazine Recognition Test (MRT; Stanovich & West, 1989), both of which have been shown to measure general exposure to print (Allen et al., 1992). Both these instruments are validated checklists based on signal detection theory that have simple cognitive requirements and reduce the tendency of participants to supply socially desirable responses about their literacy activities (West, Stanovich, & Mitchell, 1993). Siddiqui et al. created a composite index of general print exposure using their participants' performance on both the ART and the MRT. They also administered a shortened version of the Nelson–Denny Reading Comprehension subtest (Brown, Bennett, & Hanna, 1981) and a syllogistic reasoning task that was modeled on the items in Markovits and Nantel (1989). Although print exposure and syllogistic reasoning were positively correlated, a more conservative hierarchical regression analysis showed that general print exposure failed to explain a significant amount of variance in syllogistic reasoning once uni-

Helena P. Osana, Einat Idan, and Guillaume W. Jabbour, Department of Education, Concordia University, Montreal, Quebec, Canada; Guy L. Lacroix, Department of Psychology, Carleton University, Ottawa, Ontario, Canada; Bradley J. Tucker, Office of the Provost and Vice-President, Academic Affairs, Concordia University.

Portions of this work were presented by Guy L. Lacroix, Helena P. Osana, Bradley J. Tucker, and Einat Idan at the meeting of the Psychonomic Society, Minneapolis, Minnesota, 2004, and by Helena P. Osana, Guy L. Lacroix, Bradley J. Tucker, Einat Idan, and Julie Lewis at the meeting of the American Educational Research Association, Montreal, Quebec, Canada, 2005. Funding was provided by a grant awarded to Helena P. Osana by the Fonds Québécois de la Recherche sur la Société et la Culture.

We are grateful to William E. Nagy for his feedback on an earlier version of this article. We also thank Julie Lewis and Valerie Morvan for their assistance with data collection and coding.

Correspondence concerning this article should be addressed to Helena P. Osana, 1455 de Maisonneuve Boulevard Ouest, Education LB-579, Montreal, Quebec H3G 1M8, Canada. E-mail: osana@education.concordia.ca

versity experience, grade-point average (GPA), and reading comprehension were entered into the regression equation. This research suggests that if there is a significant relationship between print exposure and syllogistic reasoning, it is very small (Siddiqui et al., 1998; Stanovich, 2000).

Effects of Text Structure

Despite these results, the literature in both deductive and inductive reasoning suggests that this domain of research deserves additional attention. For example, in their study of students in Grades 5, 7, 9, and 11, Means and Voss (1996) investigated the effects of grade, ability level, and world knowledge on student argument production and evaluation. One of the main findings of their investigation was that general ability (as measured by either the Stanford–Binet or the Wechsler Intelligence Scale for Children—Revised [WISC–R]) was a stronger predictor of the soundness of arguments than age, grade, or amount of prior knowledge. Germane to our work is the speculative explanation Means and Voss offered for this result. They proposed that informal reasoning, such as argumentation, should be conceptualized in relation to language structures: There are particular forms of language, such as claims, qualifiers, and counterarguments, that are present in informal reasoning. The authors speculated that, through listening and reading, "high ability individuals may acquire particular *conventions of reasoning*, that is, forms of language employed by relatively intellectual or educated people that include, but are not limited to, language structures of argumentation" (Means & Voss, 1996, p. 165).

Further relevant to this study, Reznitskaya et al. (2001) found that the involvement of fifth graders in collaborative reasoning groups allowed them to appropriate an "argument schema" that comprised abstract and transferable knowledge of the rules, structure, and social conventions of sound argumentation. In sum, although Means and Voss (1996) and Reznitskaya et al. (2001) did not directly investigate the effects of print exposure, their speculations lend credibility to the notion that the greater individuals' exposure is to discourse that includes argument and logical forms, the more likely it is that they will acquire various forms of sound reasoning. Thus, one might hypothesize that the quality of print (i.e., type of print) to which an individual is exposed, not only its quantity, is related to his or her ability to reason both inductively and deductively.

It is well known that written texts exist in a wide variety of forms, or genres, that differ with respect to their structural components and pragmatic rules (Alexander & Jetton, 2003; Grimshaw, 2003; Jackson & Kemper, 1993; Meyer, 1999; Weaver & Kintsch, 1991). Brewer (1980) argued that the structure of expository texts reflects the underlying mental processes of logical and quasi-logical thinking, which incorporate processes such as induction, classification, and comparison. Goldman and Murray (1992) claimed that logical connectors in expository text—additive, causal, adversative, and sequential terms—serve to assist the reader to infer various logical relations and departures from these relations when he or she is attempting to comprehend informative text. Scholars have also claimed that the structure of scientific text, a specific form of expository discourse, is explanatory and, as such, hinges on the notion of cause and effect (León & Peñalba, 2002; Trabasso & Magliano, 1996).

Park and Han (2002) illustrated the deductive form of scientific explanation more explicitly. By borrowing principles from philosophy of science (e.g., Hempel, 1965), they argued that scientific discourse, whether in the form of written or verbal explanations, follows the pattern of deductive logic, in which the reason an event occurs follows logically from general laws and initial conditions that act as premises in a classical syllogism. Similarly, Reif and Larkin (1991) proposed that the processes of achieving understanding in science involve the ability to (a) create explanatory inferences that show how observable phenomena can be deduced from basic theoretical premises and (b) to reason from premises using inference rules of deductive logic.

One of the primary hypotheses for the present study is that the more individuals read expository text, and popularized scientific texts in particular, the greater is the likelihood that they will appropriate conventions and rules of sound deductive reasoning. There exists empirical evidence to suggest that individuals extract different types of information and acquire specific knowledge structures depending on the genre of the text (Chambliss, 1995; Salisbury-Glennon & Stevens, 1999; Zwaan, 1993). Chambliss (1995), for example, found that the structure of expository texts containing lengthy written arguments impacted high school students' ability to comprehend text from the point of view of argumentative reasoning. Her results indicated that students who were exposed to texts that were structured according to Toulmin's (1958) model of argument were better able to identify key reasoning structures in the text compared to students who read the same content in a topical-net format. Therefore, the results of Chambliss's (1995) study provide empirical evidence that individuals are able to assimilate normative modes of thinking when the reasoning structures are explicitly present in written text.

On the basis of the evidence presented above, therefore, we propose that a measure that is sensitive to the exposure of a particular type of print (i.e., expository texts that contain well-articulated causal and logical structures) might be a better predictor of syllogistic reasoning than an instrument that measures experiences with a wider variety of print type, including narrative, much of which may contain less than ideal models of logic and argument. In particular, we hypothesize that scientific texts written for lay audiences (e.g., the works of Stephen Jay Gould or Howard Gardner) are excellent examples of expository texts that contain explicitly delineated reasoning structures, most notably explicit statements that link events with their causes and consequences and warrants that link evidence with scientific claims, which, in turn, serve to assist the reader in building new knowledge about the domain in question (G. Myers, 1991). We therefore predict that a measure that is sensitive to print exposure quality (i.e., the exposure to scientific texts written for lay audiences) will explain more variance in syllogistic reasoning than one that measures only quantity of reading.

Effects of Inferential Processing During Reading

The research on text comprehension provides another theoretical context for the present study. It is now well known that understanding a text involves the construction of a variety of inferences that serve to establish local and global coherence among text ideas (Goldman & Rakestraw, 2000; Graesser & Bower, 1990; Graesser, Singer, & Trabasso, 1994; Trabasso & Magliano, 1996;

Zwaan & Singer, 2003). In fact, Trabasso and Magliano (1996) and Suh (1989) found that a high proportion (anywhere from 70% to 81%) of online utterances made while reading are inferences.

There is considerable variability in the number and type of inferences readers draw while reading (e.g., Graesser et al., 1994; Kintsch, 1993). For the purposes of the present study, we focus on explanatory bridging inferences that link text constituents. Zwaan and Singer (2003) defined bridging inferences as constructions made by the reader to link the current text to previous text, preserving local text coherence and thereby enhancing comprehension (e.g., J. L. Myers, Shinjo, & Duffy, 1987; Potts, Keenan, & Golding, 1988; Singer & Ferreira, 1983). In several cases, the inference involves making a causal connection between the current action, event, or state in the text and content explicitly stated earlier in the text (Graesser et al., 1994; Zwaan & Singer, 2003).

Singer, Revlin, and Halldorson (1990) likened the process of constructing explanatory bridging inferences to creating and solving mental syllogisms. As an illustration, consider the following two sentences, presented by Singer et al.: "Sue is a surgeon. Mary is a doctor, too." First, to understand how the second sentence follows from the first, one must construct the bridging inference "Sue is a doctor" (p. 36). Singer et al. proposed that to verify the acceptability of this bridge, the reader must, in effect, solve an enthymeme consisting of "Sue is a surgeon" as the first premise and the bridge ("Sue is a doctor") as the conclusion. The reader must use his or her knowledge to generate the universal statement "A surgeon is a doctor" as the missing premise, thereby accepting the bridge as reasonable and understanding how "Mary is a doctor, too" follows from "Sue is a surgeon."

Lea (1995) presented empirical evidence showing that individuals spontaneously draw logical conclusions using mental-logic principles at the moment two premises are available in text. In one part of his study, he presented short texts to undergraduate students; half of the texts contained both statements "if p then q" and "p" (inference version), and the other half contained only the first premise, "if p then q" (no-inference version). The latter prevented the participants from making the logical inference using *modus ponens.* The results demonstrated that lexical-decision targets were identified significantly faster when they followed the inference versions of texts compared to the no-inference versions, suggesting that when participants encountered the two necessary premises during reading, they spontaneously drew valid logical inferences even though these inferences were not required for comprehension. When he manipulated the texts using the *or-elimination* principle, Lea (1995) arrived at the same results.

Given the connection between constructing explanatory bridging inferences and the deductive reasoning processes demonstrated by Lea (1995; see also Lea, O'Brien, Fisch, Noveck, & Braine, 1990), we speculated that individuals who are more skilled at "filling in the gaps" (Kemper, Estill, Otalvaro, & Schadler, 1985) when trying to understand the causes and consequences of the actions in a text may be more proficient at solving deductive reasoning problems, such as syllogisms. To illustrate, we present in Table 1 two versions of "Rocky Raccoon," a text adapted from Kemper (1988). In the text in the left column, the low-inference version, all causes and consequences are explicitly linked to the actions and states described in the passage. In contrast, presented in the right column is the high-inference version of the same text,

Table 1
Excerpts From High- and Low-Inference Texts With Sample Comprehension Questions and Acceptable Responses Taken From Coding Rubric

Low-inference text	High-inference text
That day her father built a new fence around the garbage cans.	That day her father built a new fence around the garbage cans.
She wondered if Rocky would be able to get to them.	She wondered if Rocky would be able to get to them.
She did not have to wait long to see.	She did not have to wait long to see.
Very soon, Rocky ran around the back of the house	Very soon, Rocky ran around the back of the house
to where the garbage cans were,	—
and went straight to the new fence.	and went straight to the new fence.
He was surprised.	—
He sniffed.	He sniffed.
Next he scratched at the wooden boards	—
trying with his four limbs to climb up.	He tried—to climb up.
There was nothing for his curved toes to grab onto.	There was nothing for his curved toes to grab onto.
He was discouraged.	He was discouraged.
He left.	He left.

Sample questions (acceptable responses)

Why did Elizabeth's father build a fence that day? (To keep Rocky out of the garage)
Why did Rocky seem surprised? (There was a fence; he was not expecting a fence)
Why was Rocky discouraged? (He could not get to the garbage; he could not get over the fence)

Note. Story is adapted from "Inferential Complexity and the Readability of Texts," by S. Kemper, 1988, in *Linguistic Complexity and Text Comprehension: Readability Issues Reconsidered,* edited by A. Davidson & G. M. Green, Hillsdale, NJ: Erlbaum. Copyright 1988 by Erlbaum. Adapted with permission.

which is missing critical links (as indicated by dashes in Table 1) between pairs of phrases and sentences. Thus, readers who are able to comprehend critical aspects of the high-inference version do so because they must infer the missing phrases, a process that we propose is related to syllogistic reasoning.

The Present Study

The research objective in the present study was to determine whether quality of print exposure (as opposed to quantity of print only) and explanatory inference ability during reading are significant predictors of syllogistic reasoning. To our knowledge, no previous work has directly investigated the relationship between exposure to a specific genre of text and deductive reasoning. Further, few, if any, studies have used discourse processing theories to explore the link between inferential processing and syllogistic reasoning.

Our research design was closely modeled on the one used by Siddiqui et al. (1998). We presented undergraduate students with a measure of print exposure based on the ART and a parallel measure of print exposure type, developed by our research team, called the Science Masters Author Recognition Test (SMART). The SMART is based on the same signal detection theory as the ART, wherein actual popular science authors are presented among foils. We also presented participants with two measures of cognitive ability: a vocabulary checklist measure (Sá, West, & Stanovich, 1999; Stanovich et al., 1995) and Raven's Advanced Progressive Matrices (Set II; Raven, 1962). We predicted that once cognitive ability was controlled, exposure to popularized science texts would emerge as a unique predictor of syllogistic reasoning.

We used the same syllogistic reasoning task as Sá et al. (1999) for the criterion variable. This task required participants to determine whether a conclusion followed logically from two stated premises. Three types of syllogisms were included to distinguish inferential reasoning from the retrieval of factual world knowledge. Consistent syllogisms presented conclusions that were logically consistent but were also aligned with agreed-on world knowledge (e.g., rocks are not nuts). Inconsistent syllogisms provided conclusions that were not aligned with world knowledge (e.g., cigarettes are good for the health). The premises and conclusions in neutral syllogisms contained terms that were fictional (e.g., *snorlups, jamtops*). Thus, the two latter types of syllogism allowed us to assess the participants' deductive reasoning ability independently of their general world knowledge. Our measure of explanatory inference ability was adapted from the reading comprehension measure created by Kemper (1982, 1988), who used a technique based on event-chain analysis to experimentally manipulate the inference load of written texts. Using the same technique, we developed an instrument that allowed us to measure the extent to which individuals were able to create explanatory bridging inferences as they read written passages for comprehension, an aspect of reading activity that goes beyond general print exposure.

In our study, we manipulated the inference load of four texts. Two of the texts were high-inference-load texts, for which the participants were required to create explanatory inferences for comprehension, and two were low-inference-load texts, in which all sentences with critical causal explanations were included. Kemper (1988) also compared the effects of her manipulation of inference load to the effects of the well-known Flesch (1974) readability formula. In an attempt to replicate her findings, we also manipulated the low- and high-inference text versions using two different Flesch readability grade levels.

Eight open-ended comprehension questions followed each of the four texts. Our prediction was in line with that of Siddiqui et al. (1998)[1]: The better an individual is at drawing inferences from text (i.e., at performing well on comprehension questions related to high-inference load texts), the better he or she will be at evaluating conclusions that run contrary to world knowledge. This prediction is based on the assumption that the validity of an inconsistent syllogism can only be assessed with appropriate deductive logic. In contrast, performance on consistent syllogisms could require world knowledge alone and thus be independent of the ability to generate explanatory inferences from text.

Our specific research questions were the following:

1. Is exposure to popularized scientific print (i.e., scientific texts written for the lay reader) a significant predictor of syllogistic reasoning?

2. Does the ability to generate explanatory inferences during reading explain a significant amount of variance in syllogistic reasoning?

3. Is the ability to generate explanatory inferences to comprehend text a significant predictor of performance on inconsistent and neutral syllogistic tasks (as opposed to consistent syllogistic tasks)?

Method

Participants

The participants were 112 undergraduates from a large urban university in Quebec, Canada (60 women, 52 men). They were recruited through on-campus poster advertisements and verbal announcements in several undergraduate classes. The mean age was 23.7 years ($SD = 4.6$, range $= 19-49$). The participants were enrolled in 38 different disciplines, including arts, science, business, and engineering. Forty-five participants reported being in the equivalent of their 2nd year of U.S. university study, 25 were in their 3rd year, 26 were in their 4th year, and 15 reported having completed at least 4 years of university study.[2] One participant did not report year of study. The average cumulative participant GPA reported (out of 4.3) was 3.28 ($SD = 0.54$). All participants reported English as their first language. Participants received $15 CDN as compensation.

Instruments and Measures

Print exposure. We used two measures of print exposure in this study. The first was an adaptation of the ART (Allen et al., 1992; Siddiqui et al., 1998). The test used a signal detection logic, and it consisted of a list of real authors embedded among foils

[1] Siddiqui et al. (1998) made this prediction with regard to the relationship between print exposure and syllogistic reasoning. Nevertheless, the logic of the argument presented here is the same.

[2] Quebec students attend an additional year of preuniversity training.

(names of nonauthors). The participants were asked to check all the authors they recognized but not to guess because of the presence of foils. Allen et al. (1992) found that the ART was significantly correlated with reading time as reported on daily diaries and was as predictive of verbal outcomes and achievement as the diaries were (see also Stanovich & Cunningham, 1992, for more information on the validity of the ART).

Our adaptation of the ART was based on the list retrieved from West (2000a). We made appropriate modifications to customize the instrument for use with our groups of participants. Therefore, to tailor the instrument for the cohort of undergraduates involved at the time of our study, we inserted several well-known Canadian authors and other renowned international authors. In total, there were 89 names on the list, including 48 authors and 41 foils.

The second measure of print exposure specifically targeted the participants' knowledge of the popularized science literature and was named the SMART. Our research team created the instrument following the same signal detection logic as was used to create the ART. Thus, 51 popular science authors from a variety of disciplines were presented on a list that also included 39 foils, which were all names of obscure 19th-century poets. The 90 items that composed the SMART are presented in the Appendix. The instructions were identical to those of the ART; the participants were asked to identify the names of well-known science authors but not to guess because of the presence of foils.

We calculated participants' scores on both the ART and the SMART by subtracting the proportion of correctly identified authors from the proportion of checked foils. As Stanovich (2000) clearly argued in defending the checklist methodology, we do not claim that recognizing any particular science author guarantees in-depth knowledge of that author's work or even that the participant has read the recognized author's work. The assumption, however, is that familiarity with such authors suggests that the participant has been engaging in activities pertaining to science, such as reading either science articles and books or their popularized equivalents.

Cognitive ability. We used one measure of verbal ability and one measure of nonverbal ability to assess participants' general cognitive ability. The measure of verbal ability used in this study was a checklist and foils test, which has been shown to be a reliable and valid assessment of vocabulary (Anderson & Freebody, 1983; Stanovich et al., 1995). Verbal ability has also been shown to be a valid indicator of general cognitive ability (see Matarazzo, 1972). The instrument consisted of 40 real words (e.g., *connote, litany,* and *thalamic*) and 20 pronounceable nonwords (e.g., *biplaster, fusigenic,* and *relatize*). The participants were told to check the items they knew were actual words and to avoid guessing because of the presence of nonwords. The 60 items, words and foils, were taken from Zimmerman, Broder, Shaughnessy, and Underwood (1977). We calculated participants' scores on the verbal ability measure by subtracting the proportion of checked nonwords from the proportion of correctly identified existing words.

The measure of nonverbal ability was a truncated version of Raven's Advanced Progressive Matrices (Set II). Individual items and evidence of construct validity for this measure can be found in Raven (1962). The truncated version of the test involved the administration of the 18 problems most diagnostic for an undergraduate population in a time limit of 15 min. The less diagnostic items, which have been shown to produce results at floor or near ceiling, were not included (Carpenter, Just, & Shell, 1990; Raven, Court, & Raven, 1977; Sá et al., 1999). Each problem required participants to complete a logical sequence of visual stimuli. Participants' scores on the Raven were the number of correct answers generated out of a possible total of 18 problems.

Syllogistic reasoning. The participants' syllogistic reasoning skills were assessed with the 24 problems from West (2000b). The same 24 items were used in Sá et al. (1999), and the majority were originally presented in Markovits and Nantel (1989). Each problem consisted of two premises and a conclusion, and the participants were asked to decide whether the conclusion followed logically from the premises. The participants were explicitly told to assume that the premises were true and that the validity of the conclusion had to be evaluated with the information stated in the premises only.

Eight problems were classified as inconsistent because they presented a conclusion that conflicted with world knowledge. Another eight problems were classified as consistent because they presented a conclusion that agreed with world knowledge. Finally, eight problems were classified as neutral because the content of the premises and conclusions was fictional and unrelated to world knowledge. Half of the problems in each of the inconsistent, consistent, and neutral categories were valid syllogisms; the other half were invalid. It has been shown in numerous studies that participants find it more difficult to evaluate the validity of syllogisms when conclusions violate world knowledge than when they do not (Markovits & Nantel, 1989; Sá et al., 1999; Siddiqui et al., 1998). Total syllogistic reasoning scores were the total number of correct answers out of a possible 24 points. Scores for each subsection of the test (inconsistent, consistent, and neutral) were the number of correct answers out of a possible 8.

Explanatory inference ability. Explanatory inference ability was measured with a reading comprehension test that consisted of four short texts, each followed by eight open-ended comprehension questions. Three of the texts ("Johnny and His Kite," "The Mystery Spot," and "Rocky Raccoon") were taken from Kemper (1982), Kemper et al. (1985), and Kemper (1988), respectively. The fourth text was an excerpt from the short story "The Boarding House" by James Joyce (1962).

Different methods exist for manipulating the inference load of written text. One such method was employed by McNamara (2001; McNamara, Kintsch, Butler-Songer, & Kintsch, 1996), who increased text coherence by repairing low-level as well as high-level (i.e., conceptual) gaps between and among sentences in written passages. Along the same lines, Kemper's (1982, 1983, 1988) approach to increasing coherence entails making causal and temporal links explicit by inserting them in the text. Kemper's method involves using an inferential load formula for identifying the locations in text where a reader would need to make a causal inference for comprehension and repairing the missing link. Because Kemper's method specifically targets causal inferential processing, we believed her procedure to be particularly appropriate for our purposes.

Using the procedure described in Kemper (1983, 1988), therefore, we created one low-inference version and one high-inference version of each text. We first identified all the tensed and untensed clauses in the original version of each text. Tensed clauses are phrases or complete sentences that contain verbs and are inflected for tense, whereas untensed clauses are noun and verb phrase complements. Once the list of clauses was created, each clause was labeled as either an action, a physical state, or a mental state. An action is a clause that involves a change in the state of an object or person and the activities of agents in that action. States are either observable physical states or unobservable mental states. Physical states include possession, attribution, or specification, while mental states refer to enduring traits of human beings, such as emotions, cognition, or intentions (more information on these definitions can be found in Kemper, 1982, 1983, 1988).

Using Kemper's (1982, 1983, 1988) rules of event-chain analysis, we placed each clause in a tree structure that represented the sequences of actions and states that made the story logically coherent. Using the same set of rules, we were able to identify violations in the event chain—that is, gaps in the story that increased the inference load for the reader. For example, in the original version of "Rocky Raccoon," the following two action clauses were placed one after the other: "Rocky ran around the back of the house," "and went straight to the new fence" (see Table 1). According to Kemper (1988), an action–action sequence must be mediated by either an inferred physical state or an inferred mental state. To repair this particular violation and thus reduce the inference load of the text, we placed the physical state "to where the garbage cans were" between the two action clauses.

The low-inference version of each text contained no violations of the rules of event-chain analysis (i.e., no missing gaps). The high-inference version of each text had nine strategically removed clauses that created violations of the event-chain rules: Three clauses (one each from the action, physical state, and mental state categories) were removed from the beginning of the passage, three from the middle, and three from the end of the passage. The average length of the low-inference versions was 253 words, and the average length of the high-inference versions was 203 words.

In addition, using the Flesch readability scale (Flesch, 1974), we created 2 versions of each low- and high-inference-load text. Thus, for the high-Flesch version, we increased the vocabulary level and the number of compound sentences; for the low-Flesch version, we performed the opposite revisions. The average readability score across all 8 low-Flesch texts (4 high-inference and 4 low-inference) was Grade 4.13. The average readability for the high-Flesch texts was Grade 7.84. These manipulations ultimately resulted in 4 versions of each of the four texts: high inference/low Flesch, high inference/high Flesch, low inference/low Flesch, low inference/high Flesch. All 16 versions of the texts were perfectly counterbalanced with respect to order in the test booklets. Each participant received all four texts, with each text corresponding to one of the inference/Flesch versions.

Eight open-ended questions, requiring a one- to two-sentence written response, followed each of the four stories on a separate sheet of paper. The order of the questions remained constant by story. Explicit responses to the questions were not contained in the texts themselves, regardless of inference load; thus, the participants were required to access relevant aspects of the situation models they constructed as opposed to recalling specific text content. The majority of the questions began with *why* and *how,* which have been shown to uncover a wide variety of inferences (Magliano & Graesser, 1991). Sample questions can be found in Table 1.

Two trained raters independently coded all written responses to the story questions using a rubric that was provided to them before the coding began. This process consisted of coding 896 individual responses for each story (eight questions for each of 112 participants). The coding rubric consisted of all of the comprehension questions, grouped by story, each followed by 1 or 2 responses that were considered by the research team as acceptable. Acceptable responses to the sample questions are presented in Table 1. Coding involved assigning a 1 to a correct answer and a 0 to an incorrect answer. Percentage of agreement between the two coders was 95% for "Johnny and His Kite," 97% for "The Mystery Spot," 95% for "Rocky Raccoon," and 96% for "The Boarding House." All disagreements between the raters were reviewed by other project team members, and consensus was reached through discussion. In sum, the reading task was used to obtain three explanatory inference ability measures for the analyses. These measures were (a) performance on the comprehension questions for the two high-inference texts, (b) performance on the comprehension questions for the two low-inference texts, and (c) the total score on all questions for all four texts.

Procedure

Each participant received a booklet containing all the tests described in the previous section. The tests were presented in the following order in each booklet: demographic questionnaire, SMART, ART, vocabulary, Raven, syllogistic reasoning, inference ability. The participants were individually tested in small groups (of anywhere from 3 to 10 participants) in separate sessions that lasted approximately 90 min. Each participant was given a test booklet (with accompanying answer sheets for some of the instruments). Each testing session began with a brief introduction to the testing procedure. Then, for each test, the experimenter read the instructions out loud, specified the time limit, and verified that the procedure to be followed was clear before prompting the participants to begin. The participants were not permitted to proceed to the next test until the experimenter allowed them to do so. For the inference ability test, the participants were allowed 3 min to read each story and then 4 min to answer the eight comprehension questions that followed. They were not permitted to look back at the story when answering the questions.

Results

Descriptive Statistics

Cognitive ability measures. The average score on the Raven was 8.34 ($SD = 3.67$), and the Cronbach's reliability estimate was .77. The mean score on the vocabulary test was 0.53 ($SD = 0.19$), and an average of 22.52 real words and 0.56 foils were checked. The Cronbach's alpha reliability estimate on the vocabulary measure was .87.

Print exposure measures. The mean score on the ART was 0.34 ($SD = 0.17$), and an average of 16.37 real authors (with a range of 2 to 41) and 0.37 foils were checked. The Cronbach's alpha reliability estimate for this measure was .90. For the SMART, the mean score was 0.06 ($SD = 0.06$), and an average of 3.92 science authors (with a range of 0 to 19) and 0.68 foils were checked. The Cronbach's alpha reliability estimate was .81.

Explanatory inference ability. The percentage of questions answered correctly on the inference ability test was 61% ($SD =$

15%). The Cronbach's alpha reliability estimate for the test was .77. Cronbach's alpha was .65 for the low-inference-load texts and .66 for the high-inference texts, indicating comparable reliability across both inference ability subtests.

The participants' ability to answer questions about the texts as a function of vocabulary level and inference load was assessed with a 2×2 analysis of variance on the comprehension scores. The two within-subject factors were Flesch readability (low vs. high) and inference load (low vs. high). The data are presented in Table 2. A main effect was found for inference load, $F(1, 111) = 6.88$, $MSE = 0.38$, $p < .05$, but not for Flesch readability, $F(1, 111) = .80$, $MSE = 0.00$, $p > .05$. The interaction between these factors was significant, $F(1, 111) = 4.90$, $MSE = 0.12$, $p < .05$. The decomposition of this effect showed that the difference between low and high Flesch readability was significant when the inference load was high, $F(1, 111) = 4.99$, $MSE = 0.12$, $p < .05$, but not when it was low, $F(1, 111) = 0.79$, $MSE = 0.02$, $p > .05$.

Overall, these results replicate Kemper's (1988) main finding that inference load impacted the participants' response accuracy. High-inference-load texts required participants to generate causal inferences to drive comprehension, whereas low-inference-load texts explicitly provided the readers with causal connections. Thus, participants' performances on high- versus low-inference-load texts, not their performance on the high- versus low-Flesch texts, is used in the following analyses to explore the relation between the ability to generate explanatory inferences during reading and syllogistic reasoning.

Syllogistic reasoning task. The average number of correct responses for the syllogistic reasoning task was 18.59 ($SD = 3.92$). This result is similar to that found in previous studies with the same 24 syllogisms (Sá et al., 1999). The Cronbach's alpha reliability estimate for this test was .80. For consistent items, the average number of correct responses was 6.60 ($SD = 1.30$), for inconsistent items it was 5.45 ($SD = 1.79$), and for neutral items it was 6.51 ($SD = 1.62$). The Cronbach's alpha reliability estimates for the subtests in the same order were .42, .60, and .65.

Correlations Among the Main Variables

Our first step in the analysis was to generate correlations among the participant variables, the cognitive ability measures, the print exposure measures, the explanatory inference ability measures, and syllogistic reasoning ability. The results are presented in Table 3. The first two variables, university experience and GPA, were those that were used in Siddiqui et al.'s (1998) study to assess the participants' general ability and level of education. University experience was not significantly correlated with any of the syllogistic reasoning scores, replicating Siddiqui et al. (1998), nor was it correlated with the print exposure measures or vocabulary. In contrast, GPA was significantly correlated with the syllogistic reasoning measures, the inference ability measures, and vocabulary measure (but not with the Raven). Overall, therefore, these correlations show that GPA was an adequate measure of general ability and was more strongly related to other variables, including syllogistic reasoning ability, than was university experience.

The two cognitive ability measures were vocabulary and the shortened version of the Raven. The correlation between vocabulary and total syllogistic reasoning ability was significant. Vocabulary also showed the typical correlation with the ART, which replicates the findings of previous experiments (e.g., West et al., 1993), and it showed a smaller but significant correlation with the SMART ($r = .46$, $p < .001$). The difference between these two correlations was significant, $t(109) = 2.53$, $p = .01$ (see Cohen & Cohen, 1983, on comparing correlation coefficients taken from the same sample). Finally, vocabulary correlated strongly with all the inference ability measures. The correlation between the Raven and the total syllogistic reasoning score was not significant, although the Raven correlated significantly with neutral and inconsistent syllogisms. In addition, the Raven was not significantly correlated with either of the print exposure measures, but it did show moderate correlations with all of the explanatory inference ability measures. In sum, these correlations show that vocabulary was the cognitive ability variable that was the most strongly associated with syllogistic reasoning, outperforming both the Raven and GPA.

We turn now to the print exposure measures. The ART was significantly correlated with all the syllogistic reasoning scores and was also significantly correlated with all explanatory inference ability measures. In contrast, although the SMART was also significantly correlated with all syllogistic reasoning measures, it was not correlated with any of the inference ability measures. This last result makes it improbable that the SMART taps general exposure to print. Finally, the correlation between the ART and the SMART was significant ($r = .39$, $p < .001$), suggesting some degree of construct validity for the SMART. In addition, the correlations between the print exposure measures and the other variables under investigation showed that although the ART and SMART were related, the SMART was not simply an alternative measure of general exposure to print. In contrast, Siddiqui et al. (1998) used two tests to measure print exposure, the ART and the MRT, in which participants are asked to identify real magazine titles while avoiding foils. The authors found a correlation of .75 between these two tests, suggesting that the ART and MRT tap the same construct. Altogether, therefore, our results appear to indicate that the ART and the SMART, although correlated, were likely measuring different constructs.

The final set of variables to be related to syllogistic reasoning was explanatory inference ability. First, the total explanatory inference ability score was significantly correlated with the total syllogistic reasoning measure ($r = .45$, $p < .001$). The magnitude of this correlation was similar to that of the correlation between syllogistic reasoning and the Nelson–Denny Reading Comprehension test reported in Siddiqui et al. (1998),[3] which allows us to be

Table 2
Percentage of Correct Answers (with Standard Errors) Obtained on the Explanatory Inference Ability Test Comparing Flesch Readability (Low vs. High) and Inference Load (Low vs. High)

	Inference load	
Flesch Readability	Low	High
Low	61.9 (.02)	66.5 (.02)
High	59.4 (.02)	57.5 (.02)

[3] The correlation reported in Siddiqui et al. (1998) involved the same measure of syllogistic reasoning as in this study, but reading was assessed with a shortened version of the Nelson–Denny Reading Comprehension subtest (Form F; Brown et al., 1981).

Table 3
Intercorrelations Among Main Variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12
Participant variables												
1. University experience	—											
2. GPA	−.15	—										
Cognitive ability measures												
3. Vocabulary	.07	.29	—									
4. Raven	−.27	.00	.08	—								
Print exposure measures												
5. ART	.01	.13	.66	−.10	—							
6. SMART	.00	.34	.46	.09	.39	—						
Inference ability measures												
7. Low-inference texts	−.14	.16	.41	.24	.24	.06	—					
8. High-inference texts	−.10	.17	.28	.23	.20	.14	.24	—				
9. Total inference ability score	−.16	.21	.43	.30	.27	.13	.76	.81	—			
Syllogistic reasoning ability												
10. Consistent syllogisms	−.02	.20	.23	.03	.21	.19	.16	.23	.25	—		
11. Neutral syllogisms	.00	.26	.42	.16	.35	.33	.24	.43	.43	.63	—	
12. Inconsistent syllogisms	.00	.29	.45	.17	.38	.32	.22	.41	.41	.48	.57	—
13. Total syllogisms	−.01	.31	.45	.15	.39	.35	.26	.44	.45	.80	.87	.84

Note. Correlations greater than .15 are significant at the .05 level, two-tailed; correlations greater than .22 are significant at the .01 level, two-tailed; and correlations greater than or equal to .30 are significant at the .001 level, two-tailed. GPA = grade-point average; ART = Author Recognition Test; SMART = Science Masters Author Recognition Test.

confident about the validity of the inference ability measure used in this study.

To obtain a better understanding of the relationship between inferential processing and syllogistic reasoning, we also examined the relationship between syllogistic reasoning and performance on the comprehension questions for both high- and low-inference-load texts. With the exception of consistent syllogisms, the correlations between the high-inference-load measure and the syllogistic reasoning measures were greater than those found between the low-inference-load measure and the syllogistic reasoning measures (all *p*s $< .05$, as shown by *t* tests; see Cohen & Cohen, 1983). These last results appear to suggest that the ability to answer questions about texts with a high inference load is more strongly correlated with syllogistic reasoning than is the ability to answer questions about texts with a low inference load.

In sum, the correlational analyses show that vocabulary, general print exposure (ART), exposure to scientific print (SMART), explanatory inference ability, and, to a lesser extent GPA, were each related to syllogistic reasoning ability. Because these variables were intercorrelated, however, these analyses do not allow us to confirm our hypothesis that different measures of explanatory inference ability and exposure to scientific print are related to syllogistic reasoning independently of cognitive ability and general print exposure. We thus rely on the following regression analyses for this purpose.

Regression Analyses

ART and inferential ability as predictors of syllogistic reasoning. To obtain a more conservative test of the relationship between inferential processing and syllogistic reasoning, we conducted two hierarchical multiple regression analyses (following a design similar to that proposed in Siddiqui et al., 1998). The results for the first analysis are shown in Table 4 and those for the second analysis are presented in Table 5. In each regression analysis, the criterion was the total syllogistic reasoning score. The predictor variables entered in the first and second steps in each analysis were nonverbal ability (Raven) and verbal ability (vocabulary), respectively. They were chosen because in conjunction they constitute a reliable measure of general intelligence (Sá et al., 1999), and they were entered separately because the data revealed that vocabulary was more strongly correlated with syllogistic reasoning than the Raven. Hence, this order of entry allowed an independent assessment of the variance explained for both variables. GPA was not included.[4] The ART score was entered in the third step. This step allowed us to assess the variance explained by general print exposure. Finally, a different measure of explanatory inference ability was entered in the last step in each regression: low inference load in the first analysis (Table 4), and high inference load in the second analysis (Table 5).

As could be anticipated from our analysis of the correlations, the Raven did not explain much variance in total syllogistic reasoning ability. The vocabulary test did, however, accounting for a significant 20% increase in variance explained ($p < .001$). The entry of the ART in the third step did not significantly

[4] All regression analyses described in this study were also conducted with GPA entered as a predictor in a third step after the Raven and vocabulary and before the print exposure and explanatory inference ability measures. When the total syllogistic reasoning ability score was used as the criterion, GPA accounted for significant changes in the amount of variance explained, but when consistent, inconsistent, or neutral syllogisms were used, it did not. Nevertheless, the inclusion of GPA did not change the results with regard to the predictive value of our inference ability measures or of the SMART in any of the regression analyses.

Table 4
Results From Hierarchical Regression Analysis of Total Syllogistic Reasoning Score by Raven, Vocabulary, ART, and Low-Inference-Load Scores

Variable	*R*	R^2 change	*F* change	Final β	Final *F*
1. Raven[a]	.15	.02	2.39	.13	2.40
2. Vocabulary[b]	.47	.20	26.73***	.29	14.84***
3. ART[c]	.49	.02	2.88	.196	11.03***
4. Low-inference-load texts[d]	.49	.00	0.28	.05	8.29***

Note. ART = Author Recognition Test. Two participants did not provide data on the Raven.
*** $p < .001$.
[a] $df = 1, 108$. [b] $df = 2, 107$. [c] $df = 3, 106$. [d] $df = 4, 105$.

Table 6
Results From Hierarchical Regression Analysis of Total Syllogistic Reasoning Score by Raven, Vocabulary, SMART, and Low-Inference-Load Scores

Variable	*R*	R^2 change	*F* change	Final β	Final *F*
1. Raven[a]	.15	.02	2.39	.09	2.40
2. Vocabulary[b]	.47	.20	26.73***	.33	14.84***
3. SMART[c]	.49	.02	3.00	.18	11.08***
4. Low-inference-load texts[d]	.49	.01	0.73	.08	8.47***

Note. SMART = Science Masters Author Recognition Test.
*** $p < .001$.
[a] $df = 1, 108$. [b] $df = 2, 107$. [c] $df = 3, 106$. [d] $df = 4, 105$.

increase the amount of variance that the model explained, accounting for only 2% of the variance ($p > .05$).[5]

With respect to the final step in each of the analyses (low-inference-load scores in the first analysis and high-inference-load scores in the second), participants' ability to comprehend high-inference-load texts was a strong predictor of syllogistic reasoning ability. It accounted for an additional 9% of the variance ($p < .001$) in the model after nonverbal ability, verbal ability, and general print exposure had been accounted for. By contrast, participants' ability to comprehend low-inference-load text did not predict any variance (0%, $p > .05$) in syllogistic reasoning. This strongly suggests the hypothesized relationship between inference load—the ability to infer causal connections that link the sentences in texts—and the ability to assess the validity of syllogisms.

SMART as a predictor of syllogistic reasoning. To test our hypothesis concerning the possible relationship between exposure to popularized scientific print and syllogistic reasoning, we conducted two additional hierarchical multiple regression analyses. These analyses were identical to those reported in the preceding section except that the SMART was entered in place of the ART in the third step of the regression analyses. The results for the first analysis are shown in Table 6, and the results for the second are presented in Table 7. These analyses fail to provide straightforward support for our hypothesis that participants who read more scientific texts will also show better syllogistic reasoning ability. The SMART explained a nonsignificant 2% of the variance ($p > .05$) once nonverbal and verbal ability had been accounted for, which is almost identical to the variance explained by the ART. The entry of the inference ability variables in the last step of each analysis produced the same pattern of results as when the ART was used.

Despite the results of the previous regression analyses, we decided to explore the possibility that a combined measure of general and scientific print exposure might be more predictive of syllogistic reasoning ability than either the ART or the SMART individually. The logic underlying this exploration is that only people who read sufficient amounts of print, including inference-rich texts such as writings in science, will show superior syllogistic reasoning ability. Thus, a combined measure of the ART and the SMART was generated. We created and then summed *Z* scores for both these measures to create the ART/SMART composite score. We subsequently conducted two additional hierarchical regression analyses, the results of which are presented in Tables 8 and 9. The

Table 5
Results From Hierarchical Regression Analysis of Total Syllogistic Reasoning Score by Raven, Vocabulary, ART, and High-Inference-Load Scores

Variable	*R*	R^2 change	*F* change	Final β	Final *F*
1. Raven[a]	.15	.02	2.39	.07	2.40
2. Vocabulary[b]	.47	.20	26.73***	.25	14.84***
3. ART[c]	.49	.02	2.88	.17	11.03***
4. High-inference-load texts[d]	.57	.09	13.72***	.32	12.70***

Note. ART = Author Recognition Test.
*** $p < .001$.
[a] $df = 1, 108$. [b] $df = 2, 107$. [c] $df = 3, 106$. [d] $df = 4, 105$.

[5] With the present data set, we cannot fully replicate the two regression analyses presented in Siddiqui et al.'s (1998, Table 4, p. 91) article involving syllogistic reasoning and print exposure. In their hierarchical multiple regression analyses, syllogistic reasoning was the criterion, and the predictors were, in order, (a) university experience, GPA, and a print exposure measure that summed standardized scores for the ART and MRT (ARTMRTZ) for the first regression analysis and (b) university experience, GPA, the Nelson–Denny Reading Comprehension subtest (see Footnote 3), and ARTMRTZ for the second regression analysis. As Siddiqui et al. (1998) noted, the latter analysis provided the strongest test of print exposure's value as a predictor of syllogistic reasoning ability. The closest approximation with our data involves conducting a hierarchical multiple regression analysis with total syllogistic reasoning ability used as the criterion and university experience, GPA, total inference ability, and ART entered step by step in order as the predictors. The final model for this regression analysis was significant ($R = .56$), $F(105) = 11.71$, $p < .001$, and the three predictors accounted for a significant change in the variance explained: GPA ($\Delta R^2 =. 10$, $p <.001$), total inference ability ($\Delta R^2 = .15$, $p < .001$), and ART ($\Delta R^2 = .07$, $p < .01$). Thus, at face value, this analysis seems to provide support for the hypothesis that print exposure does predict syllogistic reasoning independently of university experience, general ability, and explanatory inference ability, contrary to Siddiqui et al.'s (1998) conclusion. The ART did not predict a significant amount of variance, however, in any regression analysis conducted with our data (e.g., the ones presented in Tables 4 through 9) when vocabulary was entered in a previous step. Thus, Stanovich's (2000) contention that general print exposure alone is not a very strong predictor of syllogistic reasoning ability is also supported by our data.

Table 7
Results From Hierarchical Regression Analysis of Total Syllogistic Reasoning Score by Raven, Vocabulary, SMART, and High-Inference-Load Scores

Variable	R	R^2 change	F change	Final β	Final F
1. Raven[a]	.15	.02	2.39	.03	2.40
2. Vocabulary[b]	.47	.20	26.73***	.29	14.84***
3. SMART[c]	.49	.02	3.00	.17	11.08***
4. High-inference-load texts[d]	.58	.09	14.84***	.33	13.11***

Note. SMART = Science Masters Author Recognition Test.
*** $p < .001$.
[a] $df = 1, 108$. [b] $df = 2, 107$. [c] $df = 3, 106$. [d] $df = 4, 105$.

Table 9
Results From Hierarchical Regression Analysis of Total Syllogistic Reasoning Score by Raven, Vocabulary, ART/SMART, and High-Inference-Load Scores

Variable	R	R^2 change	F change	Final β	Final F
1. Raven[a]	.15	.02	2.39	.06	2.40
2. Vocabulary[b]	.47	.20	26.73***	.20	14.84***
3. ART/SMART[c]	.50	.04	5.20*	.24	12.02***
4. High-inference-load texts[d]	.59	.09	14.28***	.32	13.71***

Note. ART/SMART = composite of the Author Recognition Test and Science Masters Author Recognition Test.
* $p < .05$. *** $p < .001$.
[a] $df = 1, 108$. [b] $df = 2, 107$. [c] $df = 3, 106$. [d] $df = 4, 105$.

analyses were identical to the previously discussed regression analyses except that the ART/SMART composite score was entered in the third step of each analysis. As before, the low-inference-load measure was entered in the final step of the first of these analyses (Table 8), and the high-inference-load measure was entered in the final step of the second (Table 9).

It can be observed that the ART/SMART composite score explained a significant 4% of the variance ($p < .05$) even when the variance related to the Raven and the vocabulary measures was accounted for. Once more, high-inference-load texts were significantly predictive of syllogistic reasoning ability ($p < .001$), whereas low-inference-load texts were not ($p > .05$). It is interesting that the levels of variance explained by these last two variables remained unchanged when explanatory inference ability measure was entered third and the ART/SMART composite was entered fourth. Hence, a conservative regression analysis controlling for verbal and nonverbal ability, and even inference ability, revealed that the combination of general print exposure and exposure to scientific print was linked to syllogistic reasoning.

The relationship between high-inference load and syllogism type. To generate further support for the hypothesis that the inferential processes used in text comprehension are related to those used in evaluating the validity of syllogisms, we conducted three final hierarchical multiple regression analyses. The results of these analyses are presented in Table 10. All regression analyses involved the same four predictor variables entered in the following order: Raven, vocabulary, the ART/SMART composite, and the high-inference-load measure of inference ability. The three measures of syllogistic reasoning—consistent, neutral, and inconsistent—served in turn as the criterion variable. If both the inference ability and the syllogistic reasoning measures involve similar inferential processes, we can predict that the relationship between high-inference-load texts and inconsistent syllogisms should be stronger than that between high-inference-load texts and consistent syllogisms. This prediction is based on the argument presented by

Table 8
Results From Hierarchical Regression Analysis of Total Syllogistic Reasoning Score by Raven, Vocabulary, ART/SMART, and Low-Inference-Load Scores

Variable	R	R^2 change	F change	Final β	Final F
1. Raven[a]	.15	.02	2.39	.11	2.40
2. Vocabulary[b]	.47	.20	26.73***	.23	14.84***
3. ART/SMART[c]	.50	.04	5.20*	.27	12.02***
4. Low-inference-load texts[d]	.51	.01	0.71	.08	9.17***

Note. ART/SMART = composite of the Author Recognition Test and Science Masters Author Recognition Test.
* $p < .05$. *** $p < .001$.
[a] $df = 1, 108$. [b] $df = 2, 107$. [c] $df = 3, 106$. [d] $df = 4, 105$.

Table 10
Results From Hierarchical Regression Analysis of Syllogism Type (Consistent, Neutral, and Inconsistent) by Raven, Vocabulary, ART/SMART, and High-Inference-Load Scores

Variable	R	R^2 change	F change	Final β	Final F
Criterion is the performance on the consistent syllogisms					
1. Raven[a]	.03	.00	0.11	−.01	0.11
2. Vocabulary[b]	.22	.05	5.36*	.08	2.70
3. ART/SMART composite[c]	.25	.01	1.54	.15	2.35
4. High-inference-load texts[d]	.30	.03	2.95	.17	2.53*
Criterion is the performance on the neutral syllogisms					
1. Raven[a]	.16	.03	2.95	.07	2.95
2. Vocabulary[b]	.44	.17	22.38***	.18	12.96***
3. ART/SMART composite[c]	.48	.03	4.22*	.22	10.30***
4. High-inference-load texts[d]	.56	.09	14.24***	.32	12.25***
Criterion is the performance on the inconsistent syllogisms					
1. Raven[a]	.17	.03	3.07	.08	3.06
2. Vocabulary[b]	.47	.19	26.43***	.22	15.11***
3. ART/SMART composite[c]	.50	.03	4.40*	.22	11.86***
4. High-inference-load texts[d]	.57	.07	11.53***	.29	12.67***

Note. ART/SMART = composite of the Author Recognition Test and Science Masters Author Recognition Test.
* $p < .05$. *** $p < .001$.
[a] $df = 1, 108$. [b] $df = 2, 107$. [c] $df = 3, 106$. [d] $df = 4, 105$.

Siddiqui et al. (1998) that the link between print exposure and reasoning should be strongest when the content of the conclusion is inconsistent with real-world knowledge and its validity can only be evaluated with deductive logic.

The expected results were found. For inconsistent syllogisms, the high-inference-load texts accounted for a significant 7% increase in the variance explained ($p < .001$), with the variance related to the Raven, vocabulary, and the ART/SMART composite controlled. The other variables that predicted a significant amount of variance were vocabulary, with 19% of the variance explained ($p < .001$), and the ART/SMART composite, which explained 3% of the variance ($p < .05$). The results for the consistent syllogisms were strikingly different. The Raven, ART/SMART, and high-inference-load texts all failed to account for a significant amount of the variance (all ps $> .05$). Vocabulary was the only variable to predict a significant 5% of the variance ($p < .05$), but this amount of variance explained is much smaller than for inconsistent syllogisms. Finally, when the score for neutral syllogisms was used as the criterion, the pattern of results was similar to that found for inconsistent syllogisms. Vocabulary ($\Delta R^2 = .17$, $p < .001$), high-inference-load texts ($\Delta R^2 = .09$, $p < .001$), and the ART/SMART composite ($\Delta R^2 = .03$, $p < .05$) each accounted for significant changes in variance when they were entered into the regression equation, while the Raven did not. This result is consistent with our hypothesis that participants who successfully comprehended the high-inference-load texts were better able to use their deductive reasoning skills as opposed to world knowledge to solve inconsistent and neutral syllogisms.[6]

Discussion

The goal of this study was to extend the work of Siddiqui et al. (1998) examining the link between print exposure and syllogistic reasoning. Motivated by recent text structure and text-processing theories (e.g., Dickson, Simmons, & Kame'enui, 1998; Graesser, Gernsbacher, & Goldman, 2003; Kintsch, 1994), we first hypothesized that exposure to certain types of print, such as popularized scientific text, would explain variance in syllogistic reasoning performance. We also hypothesized that the cognitive processes used to construct meaning from text (i.e., explanatory bridging inferences) would be predictive of the ability to reason about syllogisms. Thus, we presented 112 undergraduates with the SMART, a checklist instrument designed to measure exposure to popularized scientific literature; the ART, a validated measure of general print exposure; an explanatory inference ability test that measured participants' facility with comprehending texts with high and low inference loads; and a syllogistic reasoning test that measured the ability to reason through logical problems that were either consistent, inconsistent, or neutral with respect to world knowledge.

Replicating the finding of Siddiqui et al. (1998), we found a significant zero-order correlation between general print exposure (ART) and syllogistic reasoning. More important, however, once we entered the ART into a hierarchical regression analysis following the nonverbal and verbal measures of general ability (Raven and vocabulary, respectively), the ART no longer significantly predicted syllogistic reasoning performance. This result is similar to that of Siddiqui et al., who concluded that their composite measure of general print exposure (a combination of the ART and the MRT) did not explain a significant amount of variance in syllogistic reasoning once general ability and reading comprehension were controlled in a hierarchical regression analysis. Our measure of scientific print exposure (SMART) yielded similar results. When the SMART was entered as the third step in the regression equation instead of the ART, it, too, failed to explain a significant amount of variance in syllogistic reasoning.

One of our most salient findings occurred, however, when we combined the ART and the SMART into a composite score. The significant amount of variance explained by the ART/SMART composite (4%) is striking given that it is independent of the variance associated with vocabulary and comprehension of high-inference-load texts (together, almost 30%). This effect is all the more interesting because only 3.95 science authors were recognized on the SMART (cf. 16.66 authors on the ART), which indicates that many undergraduates in our sample had limited exposure to popularized scientific literature. Thus, one could have expected the SMART to be of little diagnostic value. Yet the zero-order correlations between the ART and syllogistic reasoning and between the SMART and syllogistic reasoning were similar. In addition, the moderate correlation between the ART and the SMART showed that the tests were not simply interchangeable measures of general print exposure. Thus, we suspect that the SMART was successful in capturing the participants' exposure to popularized scientific print and, consequently, the logical structures and causal arguments presented therein.

Nevertheless, because of the participants' limited interactions with this type of print, the exposure to scientific literature in itself was insufficient to predict any variance in syllogistic reasoning ability once general cognitive variables were taken into account. When we added the general print exposure measure to the SMART, we obtained a composite measure that was successful in detecting an association between exposure to print and syllogistic reasoning not mediated by other cognitive variables. We believe that the ART/SMART was more sensitive to syllogistic reasoning precisely because it captured some exposure to logical structure, though in a less concentrated form because of the variety of genres exemplified by the authors it contains. The data also suggest that there may be a SMART cutoff, below which the effects of such exposure are minimal. Further study of this hypothesis is warranted by our findings.

Our results clearly highlight the methodological challenge of efficiently quantifying an individual's exposure to texts that contain well-articulated causal and logical structures. The argument made by various reading scholars (e.g., Stanovich & Cunningham, 1992) linking general print exposure and vocabulary is straightforward. The vocabulary encountered in print is typically much richer than that encountered in speech (Hayes & Ahrens, 1988); thus, people who engage in a great deal of reading activity, no matter the quality of print, should develop a better vocabulary than those who do not. In effect, empirical research, including the present study, has shown that print exposure and vocabulary are

[6] The different pattern of results for the consistent syllogisms could also be attributed to the relatively low internal reliability of the consistent syllogisms subtest (Cronbach's $\alpha = .42$) compared to the reliability of the inconsistent and neutral subtests (Cronbach's αs of .60 and .65, respectively).

strongly related (Cunningham & Stanovich, 1991; Stanovich, 2000). As discussed in the introduction, however, not all text genres expose people to the same quantity and quality of logical arguments (e.g., Graesser, McNamara, & Louwerse, 2003; Meyer, 1999; Weaver & Kintsch, 1991), which makes it particularly challenging to develop a measure that is sensitive to such exposure. Therefore, to successfully use the signal detection logic underlying recognition tests, we chose to focus on a specific literary domain, namely the popularized scientific literature, which offers, in all probability, a higher concentration of logical arguments. Thus, if a person is exposed to a large quantity of scientific print, then that person should be exposed to many logically stated arguments and should thus theoretically also show better syllogistic reasoning skills (or so we hypothesized).

It does not follow, however, that if a person is exposed to many logically stated arguments, then that person reads large quantities of scientific print. Other literary domains, such as philosophy, for example, may also be abundant in causal and logical structures. Hence, the SMART is limited in that it can only tap a subset of all print exposure potentially conducive to developing sound deductive reasoning skills. On a related point, it is also possible that an individual's education (i.e., field of study) or background experience, such as exposure to a logic course, may explain differences in performance on the SMART. Notwithstanding the methodological difficulties, we believe the present results are sufficiently encouraging to pursue this line of research. A future replication with participants who, on the whole, have greater exposure to the scientific literature might allow for a stronger link between reasoning ability and print exposure as measured by the SMART. Another study in which background experience can be used as an independent variable might yield important information on the characteristics of the instrument and may also uncover other possible sources of logical forms.

The second major focus of this study was to relate the ability to construct explanatory bridging inferences with performance on syllogistic reasoning tasks. Using a measure that is more sensitive to the inferential processes used during reading than are standardized measures of reading comprehension, such as the Nelson–Denny subtest (Brown et al., 1981), we showed that the participants' ability to answer questions about high-inference-load texts was predictive of syllogistic reasoning ability, even when nonverbal reasoning, vocabulary, and general print exposure were controlled (see Radvansky & Copeland, 2004, for a related result). Under the same conditions, however, the ability to answer questions about the low-inference-load texts was not predictive of syllogistic reasoning. Moreover, subsequent regression analyses showed that the effect of high-inference-load comprehension held only when the validity of the syllogisms had to be evaluated on the basis of logic and not world knowledge; participants' ability to answer questions about high-inference-load texts was not predictive of the results for consistent syllogisms. These results clearly support our main hypothesis—namely, that the ability to produce inferences during reading comprehension is related to syllogistic reasoning.

We propose two reasons for the relationship between inference generation and syllogistic reasoning performance. First, it is possible that individuals who are skilled at constructing causal explanations when attempting to make sense of high-inference-load texts use a greater number of comprehension-monitoring devices that assist them to construct coherent models of the texts. In a related line of research, Coté and Goldman (1999) demonstrated that readers who construct coherent situation models use active and strategic monitoring processes to reinstate critical aspects of the text so that they remain in working memory long enough for causal connections to be made. Such practice in connecting pieces of information extracted from the text and reflecting on derived meanings may explain the participants' enhanced facility with deductive reasoning tasks. Another possibility is that individuals who are practiced at making causal connections when reading inference-eliciting texts have acquired specific knowledge structures, or schemas, about the conventions and rules of deductive logic. While a similar idea has been forwarded by others (Means & Voss, 1996; Reznitskaya et al., 2001), further studies are needed to directly investigate the nature of the effects of discourse processes on higher order reasoning, both deductive and inductive.

Considered globally, our results suggest that verbal ability is strongly related to syllogistic reasoning: Vocabulary and the ability to generate causal inferences in text comprehension were both strong predictors in all our analyses when the validity of the syllogisms could not be determined from general world knowledge. These results are consistent with recent theories and research suggesting that verbal ability and verbal working memory play a large role in inductive and deductive reasoning (Almor & Sloman, 2000; Capon, Handley, & Dennis, 2003; Radvansky & Copeland, 2004). Perhaps more important, our results may provide the strongest evidence to date that reading is closely tied to the development of reasoning ability. It is not difficult to believe that explanatory inference ability improves as the exposure to the number of texts increases or as the exposure to inference-eliciting structures increases. Given that the link between reading and the development of vocabulary is well established, we suggest that the impact of print exposure on reasoning ability may be indirect: Print exposure leads to development of verbal ability, which, in turn, leads to better reasoning skills. Clearly, additional research is needed to investigate more directly which mediating variables are present in the development of reasoning through text.

In sum, our data reveal two major findings: (a) Individuals who are exposed to texts with well-delineated written arguments and logical structures are more adept at solving syllogistic reasoning tasks, and (b) individuals who are more proficient at drawing connections between given pieces of information in a text are better equipped to reason logically about syllogisms. At first glance, these findings may appear contradictory. On the one hand, arguments presented in full assist readers to make appropriate syllogistic inferences, perhaps because they have appropriated effective reasoning schemas that allow them to do so. This finding confirms the conclusions reached by others (e.g., Chambliss, 1995; Means & Voss, 1996). On the other hand, our data also suggest that readers who must construct lines of argument for themselves are more proficient at solving deductive reasoning tasks. This latter finding is also consistent with another line of research demonstrating that there may be sound pedagogical reasons to leave some text with less explicit connections. In his review of the literature, Kintsch (1994) indicated that the reader's prior knowledge and the type of comprehension task are important variables to consider in evaluations of the effects of low- versus high-coherence texts. McNamara et al. (1996) found, for example, that on free recall questions, participants performed better with high-coherence texts, no matter the extent of their prior knowledge. For questions that demanded more active construction processes, such as those requiring problem solving or inference generation, the high-knowledge participants performed significantly better with low-coherence texts. Whether or not the same effect would be revealed with a deductive reasoning task as the criterion remains open.

In light of the work reviewed by Kintsch (1994), therefore, our two seemingly disparate findings may not be as contradictory as they appear. The prior knowledge of our participants could have differentiated those who benefited from exposure to low-inference texts from those who were better at constructing meaning from high-inference texts (see also Voss, Fincher-Kiefer, Greene, & Post, 1986). Furthermore, as the results of McNamara et al. (1996) indicate, the effect of prior knowledge may also interact with the type of question used on the inference ability measure, which was not a focus of the present study. Therefore, our conclusions, together with previous research, appear to suggest that students can acquire higher order reasoning skills if they are exposed to a variety of high- and low-inference-load texts, but that certain conditions may make one type of text more beneficial than the other. That is, students' prior knowledge and the types of questions asked in the classroom could contribute to their development as thinkers in important ways. The details of such learning and pedagogical mechanisms are as yet not well understood and deserve further attention from researchers.

References

Alexander, P. A., & Jetton, T. L. (2003). Learning from traditional and alternative texts: New conceptualizations for the information age. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 199–241). Mahwah, NJ: Erlbaum.

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology, 84,* 489–503.

Almor, A., & Sloman, S. A. (2000). Reasoning versus memory in the Wason selection task: A non-deontic perspective on perspective effects. *Memory & Cognition, 28,* 1060–1069.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Huston (Ed.), *Advances in reading/language research* (Vol. 2, pp. 231–256). Greenwich, CT: JAI Press.

Brewer, W. F. (1980). Literary theory, rhetoric, and stylistics: Implications for psychology. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 221–239). Hillsdale, NJ: Erlbaum.

Brown, J., Bennett, J., & Hanna, G. (1981). *The Nelson–Denny Reading Test.* Lombard, IL: Riverside.

Capon, A., Handley, S. J., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking and Reasoning, 9,* 203–244.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97,* 404–431.

Chambliss, M. J. (1995). Text cues and strategies successful readers use to construct the gist of lengthy written arguments. *Reading Research Quarterly, 30,* 778–807.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. V. Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169–193). Mahwah, NJ: Erlbaum.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83,* 264–274.

Dickson, S. V., Simmons, D. C., & Kame'enui, E. J. (1998). Text organization: Research bases. In D. C. Simmons & E. J. Kame'enui (Eds.), *What reading research tells us about children with diverse learning needs* (pp. 239–277). Mahwah, NJ: Erlbaum.

Flesch, R. (1974). *The art of readable writing* (2nd ed.). New York: Harper & Row.

Goldman, S. R., & Murray, J. D. (1992). Knowledge of connectors as cohesion devices in text: A comparative study of native-English and English-as-a-second-language speakers. *Journal of Educational Psychology, 84,* 504–519.

Goldman, S. R., & Rakestraw, J. A., Jr. (2000). Structural aspects of constructing meaning from text. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 311–335). Mahwah, NJ: Erlbaum.

Graesser, A. C., & Bower, G. H. (Eds.). (1990). *Inferences and text comprehension.* San Diego, CA: Academic Press.

Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (2003). Introduction to the handbook of discourse processes. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 1–23). Mahwah, NJ: Erlbaum.

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82–98). New York: Guilford Press.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395.

Grimshaw, A. D. (2003). Genres, registers, and contexts of discourse. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 25–82). Mahwah, NJ: Erlbaum.

Guthrie, J. T., Wigfield, A., Metsala, J. L., & Cox, K. E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading, 3,* 231–256.

Hayes, D. P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of "motherese"? *Journal of Child Language, 15,* 395–410.

Hempel, C. G. (1965). *Aspects of scientific explanation.* New York: Free Press.

Jackson, J. D., & Kemper, S. (1993). Age differences in summarizing descriptive and procedural texts. *Experimental Aging Research, 19,* 39–51.

Joyce, J. (1962). *Dubliners.* New York: Viking Press.

Kemper, S. (1982). Filling in the missing links. *Journal of Verbal Learning and Verbal Behavior, 21,* 99–107.

Kemper, S. (1983). Measuring the inference load of a text. *Journal of Educational Psychology, 75,* 391–401.

Kemper, S. (1988). Inferential complexity and the readability of texts. In A. Davidson & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 141–165). Hillsdale, NJ: Erlbaum.

Kemper, S., Estill, R., Otalvaro, N., & Schadler, M. (1985). Questions of facts and questions of inferences. In A. C. Graesser & J. B. Black (Eds.), *The psychology of questions* (pp. 227–246). Hillsdale, NJ: Erlbaum.

Kintsch, W. (1993). Information accretion and reduction in text processing: Inferences. *Discourse Processes, 16,* 193–202.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49,* 294–303.

Lea, R. B. (1995). On-line evidence for elaborative logical inferences in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1469–1482.

Lea, R. B., O'Brien, D. P., Fisch, S. M., Noveck, I. A., & Braine, M. D. S. (1990). Predicting propositional logic inferences in text comprehension. *Journal of Memory and Language, 29,* 361–387.

León, J. A., & Peñalba, G. E. (2002). Understanding causality and temporal sequence in scientific discourse. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 155–178). Mahwah, NJ: Erlbaum.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics, 20,* 193–232.

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition, 17,* 11–17.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore: Williams & Wilkins.

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, E., Butler-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.

Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction, 14,* 139–178.

Meyer, B. J. F. (1999). Importance of text structure in everyday reading. In A. Ram & K. Moorman (Eds.), *Understanding language understanding* (pp. 227–252). Cambridge, MA: MIT Press.

Myers, G. (1991). Lexical cohesion and specialized knowledge in science and popular science texts. *Discourse Processes, 14,* 1–26.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). The role of causal relatedness and memory. *Journal of Memory and Language, 4,* 453–465.

Park, J., & Han, S. (2002). Using deductive reasoning to promote the change of students' conceptions about force and motion. *International Journal of Science Education, 24,* 593–609.

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language, 27,* 399–415.

Radvansky, G. A., & Copeland, D. E. (2004). Reasoning, integration, inference alteration, and text comprehension. *Canadian Journal of Experimental Psychology, 58,* 133–141.

Raven, J. C. (1962). *Advanced Progressive Matrices (Set II*). London: H. K. Lewis.

Raven, J. C., Court, J. H., & Raven, J. (1977). *Manual for Advanced Progressive Matrices (Sets I and II).* London: H. K. Lewis.

Reif, F., & Larkin, J. H. (1991). Cognition in scientific and everyday domains: Comparison and learning implications. *Journal of Research in Science Teaching, 28,* 733–760.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S.-Y. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32,* 155–175.

Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91,* 497–510.

Salisbury-Glennon, J. D., & Stevens, R. J. (1999). Addressing preservice teachers' conceptions of motivation. *Teaching and Teacher Education, 15,* 741–752.

Sénéchal, M., LeFevre, J.-A., Thomas, E. M., & Daley, K. E. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly, 33,* 96–116.

Siddiqui, S., West, R. F., & Stanovich, K. E. (1998). The influence of print exposure on syllogistic reasoning and knowledge of mental-state verbs. *Scientific Studies of Reading, 2,* 81–96.

Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior, 22,* 437–448.

Singer, M., Revlin, R., & Halldorson, M. (1990). Bridging-inferences and enthymemes. In A. C. Graesser & G. H. Bower (Eds.), *Inferences and text comprehension* (pp. 35–51). San Diego, CA: Academic Press.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers.* New York: Guilford Press.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition, 20,* 51–68.

Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology, 85,* 211–229.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24,* 402–433.

Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology, 31,* 811–826.

Suh, S. (1989). *Causal inferences during text comprehension.* Unpublished doctoral dissertation, University of Chicago.

Toulmin, S. E. (1958). *The uses of argument.* New York: Cambridge University Press.

Trabasso, T., & Magliano, J. (1996). How do children understand what they read and what can we do to help them? In M. Graves, P. Van den Brock, & B. Taylor (Eds.), *The first R: Every child's right to read* (pp. 160–188). New York: Teachers College Press.

Voss, J. F., Fincher-Kiefer, R. H., Greene, T. R., & Post, T. A. (1986). Individual differences in performance: The contrastive approach to knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 297–334). Hillsdale, NJ: Erlbaum.

Weaver, C. A., & Kintsch, W. (1991). Expository text. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 230–245). Mahwah, NJ: Erlbaum.

West, R. F. (2000a). *Author Recognition Questionnaire key—CART.HS (V3*). Retrieved April 11, 2005, from http://falcon.jmu.edu/%7Ewestrf/stimuli/cart_key.html

West, R. F. (2000b). Syllogisms 4.0. Retrieved March 2004 from http://falcon.jmu.edu/%7Ewestrf/stimuli/syllogisms_v4_key.html

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly, 28,* 34–50.

Zimmerman, J., Broder, P. K., Shaughnessy, J. J., & Underwood, B. J. (1977). A recognition test of vocabulary using signal-detection measures, and some correlates of word and nonword recognition. *Intelligence, 1,* 5–31.

Zwaan, R. A. (1993). *Aspects of literary comprehension: A cognitive approach.* Philadelphia: John Benjamins.

Zwaan, R. A., & Singer, M. (2003). Text comprehension. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 83–121). Mahwah, NJ: Erlbaum.

(*Appendix follows*)

Appendix

Science Masters Author Recognition Test

INSTRUCTIONS: *Below is a list of names. Some of the people in the list are scientists and writers of popular science books and some are not. Please read the names and place a check mark next to the names of those individuals whom you know to be scientists. Do not guess. Check only those whom you know to be scientists. Remember, some of the names on this list are people who are not scientists, so guessing can easily be detected.*

☐ 1. George Moreby Acklom
☐ 2. T. B. Aldrich
☐ 3. P. W. Atkins
☐ 4. Robert Axelrod
☐ 5. John D. Barrow
☐ 6. Katherine L. Bates
☐ 7. Gregory Bateson
☐ 8. Mary Catherine Bateson
☐ 9. Gwendolyn B. Bennet
☐ 10. William H. Calvin
☐ 11. David J. Chalmers
☐ 12. D. Coolbrith
☐ 13. Frederick C. Crews
☐ 14. Paul Davies
☐ 15. Richard Dawkins
☐ 16. Daniel C. Dennett
☐ 17. Aubrey De Vere
☐ 18. Frans de Waal
☐ 19. Jared Diamond
☐ 20. Bernard Dixon
☐ 21. Thomas Doubleday
☐ 22. Gerald Edelman
☐ 23. Ethel Edwards
☐ 24. Susan Evance
☐ 25. Richard Feynman
☐ 26. D. Ficke
☐ 27. Helen Frazee-Bower
☐ 28. Howard Gardner
☐ 29. Martin Gardner
☐ 30. Murray Gell-Mann
☐ 31. Stephen Jay Gould
☐ 32. Susan Greenfield
☐ 33. Paul R. Gross
☐ 34. Louise Guiney
☐ 35. Ian Hacking
☐ 36. William R. Hamilton
☐ 37. Stephen Hawking
☐ 38. May Herschel-Clarke
☐ 39. Daniel Hillis
☐ 40. Douglas R. Hofstadter
☐ 41. John William Inchbold
☐ 42. Thomas C. Irwin
☐ 43. Julian Jaynes
☐ 44. Sophie Jewett
☐ 45. George F. Judd
☐ 46. Frank Kermode
☐ 47. F. C. Kolbe
☐ 48. Emma Lazarus
☐ 49. Richard Leakey
☐ 50. Norman Leavitt
☐ 51. R. C. Lewontin
☐ 52. Stuart Livingstone
☐ 53. Elizabeth F. Loftus
☐ 54. Lynn Margulis
☐ 55. P. B. Medawar
☐ 56. Richard M. Milnes
☐ 57. Marvin Minsky
☐ 58. Charles Mulvany
☐ 59. R. A. Nelson
☐ 60. David Novak
☐ 61. S. James O'Grady
☐ 62. M. F. Ossoli
☐ 63. Gilbert Parker
☐ 64. Janetta Philipps
☐ 65. Steven Pinker
☐ 66. Thomas Rickmann
☐ 67. Jeremy Rifkin
☐ 68. Harriet Robinson
☐ 69. Frank Ryan
☐ 70. Oliver Sacks
☐ 71. Carl Sagan
☐ 72. W. W. Sawyer
☐ 73. Clinton Scollard
☐ 74. Stephen C. Schneider
☐ 75. John Searle
☐ 76. Robert Shapiro
☐ 77. George Smoot
☐ 78. Ian Stewart
☐ 79. Marion Strobel
☐ 80. Edward Thurlow
☐ 81. Frederick Goddard Tuckerman
☐ 82. Louis Untermeyer
☐ 83. Washington Van Dusen
☐ 84. Augusta Webster
☐ 85. Robert A. Weinberg
☐ 86. Steven Weinberg
☐ 87. Mary Wroth
☐ 88. George C. Williams
☐ 89. Edward O. Wilson
☐ 90. Danah Zohar

Received March 9, 2006
Revision received March 27, 2007
Accepted April 30, 2007 ■