

Michael Pressley

University of Maryland

Elizabeth S. Ghatala

University of Houston

Vera Woloshyn

Jennifer Pirie

University of Western Ontario

Sometimes adults miss the main ideas and do not realize it: Confidence in responses to short-answer and multiple-choice comprehension questions

IN TWO experiments Canadian university students read challenging passages, each of which was followed by a short-answer or multiple-choice question covering some content in the passage. In the first experiment, each student was asked to read a passage and answer the accompanying question, and then to make a decision whether to move forward (if he or she thought the answer was probably correct) or to look back in the text and try the question again (if he or she believed the response was probably incorrect). As found in previous research, students' monitoring of their reading and rereading was slightly better in the short-answer than in the multiple-choice condition. More striking, however, was the finding that students rarely chose to look back for general, thematic questions (as contrasted with detail questions), even when their answers were incorrect. In Experiment 2, students were asked directly to rate their confidence in their answers to short-answer and multiple-choice questions. As in the first study, students had great confidence that their answers to thematic questions in both short-answer and multiple-choice formats were correct, even when they were wrong. Importantly, students' overconfidence in answers to thematic questions was not related to their verbal ability. The authors conclude that when adults read challenging, inconsiderate texts, they may often be unaware of gross comprehension problems. Future research is necessary to determine how common such serious misperceptions are among adults.

Il arrive que les adultes n'identifient pas les idées importantes et ne s'en rendent pas compte: Niveaux de confiance à des questions de compréhension

AU COURS DE deux expériences, des étudiants de niveau universitaire devaient lire de courtes textes présentant un certain niveau de difficulté, chacun étant suivi de deux types de questions: (a) questions à court développement; (b) questions à choix multiples. Au cours de la première expérience, chaque étudiant lisait d'abord un des textes puis était invité à répondre aux questions qui suivaient; il devait ensuite décider s'il poursuivait sa lecture ou s'il relisait le texte selon qu'il avait le sentiment d'avoir compris ou non. Conformément à des résultats obtenus antérieurement, les décisions des étudiants étaient meilleures aux questions à courts développements qu'aux questions à choix multiples. De façon surprenante, les étudiants décidèrent rarement de relire le texte pour vérifier leur réponse aux questions thématiques par opposition aux questions portant sur des détails, même lorsque leurs réponses étaient fausses. Dans la seconde expérience, les étudiants devaient coter leurs réponses selon une échelle de confiance. Conformément à ce qui a été trouvé dans la première expérience, les cotes aux questions thématiques furent plus élevées que les cotes aux questions de détails même lorsqu'elles étaient fausses. Des corrélations établies entre le niveau de confiance dans les réponses aux questions thématiques et le niveau d'habileté verbale ne furent pas significatives. En conclusions, les auteurs font ressortir le fait que lorsqu'ils sont confrontés à des textes difficiles et plus ou moins bien construits, les adultes ne réalisent pas toujours qu'ils ont des problèmes importants de compréhension. Des études plus poussées seraient nécessaires pour vérifier s'il s'agit là d'un phénomène généralisé chez les adultes.

Algunas veces a los adultos se les escapan las ideas principales sin darse cuenta: La confianza en las respuestas a preguntas para evaluar la comprensión de respuesta corta y de opción múltiple

EN DOS EXPERIMENTOS llevados a cabo entre estudiantes universitarios, se pidió a estos que leyeran pasajes difíciles, seguidos inmediatamente por una pregunta de respuesta corta o una pregunta de opción múltiple que cubría cierto contenido del pasaje. En el primer experimento, se le pidió a cada estudiante que leyera primero un pasaje y contestara la pregunta correspondiente. Al terminar, se le pedía que tomara la decisión de continuar (si creía que su respuesta era correcta), o regresar en el texto e intentar contestar la misma pregunta otra vez (si pensaba que la respuesta estaba incorrecta). Como se había encontrado en investigaciones previas, los estudiantes monitoreaban su lectura mejor en la condición de respuesta corta que en la de opción múltiple. Sorprendentemente, sin embargo, los estudiantes apenas sí elegían regresar al texto para revisar en las preguntas de tema general (en contraste con preguntas detalladas), aún en los casos en que las respuestas eran incorrectas. En el experimento 2, se les pidió directamente a los estudiantes que evaluaran la confianza que sentían en sus respuestas a preguntas de respuesta corta y de opción múltiple. Como en la primera investigación, los estudiantes demostraron gran confianza en que estaban correctas sus respuestas a las preguntas temáticas, tanto para respuesta corta como de opción múltiple, aún cuando, en realidad, estuvieran incorrectas. El exceso de confianza por parte de los estudiantes en sus respuestas a las preguntas temáticas no estaba relacionado con su habilidad verbal. Los autores llegaron a la conclusión que cuando los adultos leen textos difíciles, pueden con frecuencia, dejar pasar desapercibidos problemas graves de comprensión. Se necesita más investigación que determine si tales errores en percepción son comunes entre adultos.

Manchmal erkennen Erwachsene die Kernaussagen nicht und bemerken es nicht: Vertrauen in Antworten zu Verständnisfragen, die Kurzantworten und Multiple-Choice verwenden

IN ZWEI EXPERIMENTEN lasen Studenten schwierige Textabschnitte und beantworteten anschließend jeweils eine Kurzantwort-oder Multiple-Choice-Frage, die sich mit einem Teil des Lesestücks befaßte. Im ersten Experiment wurde jeder Student gebeten, zuerst einen Abschnitt zu lesen und die begleitende Frage zu beantworten und erst danach zu entscheiden, ob er/sie weiterlesen soll (wenn er/sie glaubt, daß die Antwort richtig ist) oder im Text zurückgehen und die Beantwortung der Frage noch einmal versuchen soll (wenn er/sie glaubt, daß die Antwort wahrscheinlich falsch ist). In früheren Forschungen war bereits festgestellt worden, daß Studenten das Lesen und Nachlesen bei den Kurzantwort-Fragen etwas besser kontrollierten als bei den Multiple-Choice-Fragen. Überraschenderweise schauten Studenten jedoch kaum noch einmal nach, wenn es sich um allgemeine, thematische Fragen (im Gegensatz zu Detailfragen) handelte—auch dann, wenn die Antworten falsch waren. Im zweiten Experiment wurden Studenten gebeten, das Vertrauen in ihre Antworten zu Kurzantwort- und Multiple-Choice-Fragen unmittelbar einzustufen. Wie bereits in der ersten Studie, so zeigten Studenten auch hier ein großes Vertrauen, daß ihre Antworten zu thematischen Fragen des Kurzantwort- und Multiple-Choice-Formats richtig waren, selbst wenn sie falsch waren. Wichtig dabei ist, daß das übermäßige Vertrauen in Antworten zu thematischen Fragen nicht in bezug zu der verbalen Fähigkeit der Studenten stand. Die Verfasser schließen daraus, daß Erwachsene sich oft nicht über wesentliche Verständnisprobleme im klaren sein dürften, wenn sie schwierige, unbesonnene Texte lesen. Weitere Forschungen sind erforderlich, um festzustellen, ob solche ernsthaften Fehleinschätzungen bei Erwachsenen allgemein üblich sind.

How do readers know when they do not understand a passage? Readers can use strategies such as analyzing the structure of the text, summarizing the main ideas, and rereading parts of the material that are not understood initially (e.g., Baker & Brown, 1984) in order to increase their understanding of difficult text, but they are not likely to try to improve their comprehension unless they realize that they have not entirely understood a passage. The assumption that monitoring comprehension is necessary to the effective use of comprehension strategies has stimulated much research on comprehension monitoring, beginning with Markman's studies (e.g., 1977, 1979).

Markman pioneered the use of error detection tasks to test comprehension monitoring. In her studies, children read or listened to stories that contained internal contradictions. For example, one part of one story specified, "Fish must have light in order to see. There is abso-

lutely no light at the bottom of the ocean," but a later passage in the same story stated, "Some fish that live at the bottom of the ocean can see the color of their food; that is how they know what to eat" (Markman, 1979, p. 646). Each child was asked to change the stories in order to make them easier to understand. Markman found that children often overlooked internal contradictions; she concluded that children often fail to monitor prose for internal consistency. Findings from a number of studies conducted since Markman's initial investigations have suggested that readers' detection of inconsistencies increases with age and reading ability (see Wagoner, 1983, for a review).

However, Winograd and Johnston (1982) have contested Markman's use of error detection tasks to measure comprehension monitoring. They point out that a reader's failure to detect inconsistencies could be due to factors other than ineffective monitoring; they detail 12

alternative explanations for such a result. For example, readers might not notice errors if they lacked sufficient background knowledge. Second, readers (especially young readers) might have a general faith that writers provide unambiguous communications. Third, readers might believe that printed texts could not contain errors. Fourth, readers might make alternative interpretations of the text that would render it consistent. Fifth, readers might notice a discrepancy but believe that subsequent information in the text would resolve it.

Consequently, some investigators have tried to examine comprehension monitoring in other ways. One such method involves a performance prediction task, in which subjects are asked to estimate how well they expect to do on a test covering material that they have just read. Although the specific procedures have varied from study to study, one finding has been obtained consistently: Adults usually cannot predict accurately how well they will do on tests of comprehension and memory, even when their awareness is assessed as soon as they have finished reading a passage (e.g., Epstein, Glenberg, & Bradley, 1984; Glenberg & Epstein, 1987; Glenberg, Wilkinson, & Epstein, 1982; Maki & Berry, 1984; Pressley, Snyder, Levin, Murray, & Ghatala, 1987). In other words, either adults do not monitor their understanding and acquisition of information as they read, or they do so imperfectly.

In addition, some researchers have begun to identify conditions that can improve adults' awareness of whether text content has been learned. One such condition is to require students to answer questions as they read. For example, in an earlier study, we (Pressley et al., 1987) asked adult subjects to read a chapter of introductory psychology material. In part of that study, subjects were asked as soon as they had finished reading a passage to predict how well they would do on a 15-item test that required them to recall passage details. The accuracy of these predictions was best when students had read a version of the passage that included adjunct questions that were similar to the items that would be on the test. We concluded that, by attempting to answer adjunct questions as

they read, readers obtained information about whether they had comprehended the text well enough to be able to answer the questions on the test.

If answering adjunct questions can increase a student's awareness of whether learning from text has occurred, then answering questions might also promote appropriate regulation of rereading strategies. That is, inability to answer an adjunct question correctly should be a cue to the reader to look back and to process the text further (e.g., Garner, 1987, Chapter 3). Because looking back is one of the easiest reading comprehension strategies that students can execute, and because adjunct questions can be added to text easily, designing texts that include such questions offers great potential for encouraging readers to reinspect the text when necessary. It is important, therefore, to determine the characteristics of adjunct questions that best facilitate effective comprehension monitoring.

Some earlier studies have suggested that some question formats may lead to more appropriate regulation of rereading than other formats do. For example, in one study (Pressley & Ghatala, 1988), university students read passages from the verbal portion of the Scholastic Aptitude Test (SAT) and answered multiple-choice questions about those passages. Students often believed that there was a good chance that they had answered correctly when their answers were, in fact, incorrect. Unfortunately, students with such errant beliefs are unlikely to reread the text, even though their incorrect responses indicate that rereading the passage would improve their comprehension.

Why did students have high confidence in incorrect answers to questions in multiple-choice format? One possible explanation is that good distractors for multiple-choice questions tap information covered in the text, and the familiarity of the distractor might reinforce the perception that the distractor is the correct answer. This hypothesis is supported by basic theory and by the results of research on factors affecting performance on multiple-choice items (Ekstrand, Wallace, & Underwood, 1966; Ghatala & Levin, 1976).

Question formats other than multiple-choice, in which readers must construct their own answers to the questions or retrieve answers from the text, might produce a more accurate awareness of whether an answer is incorrect. For example, readers presented with short-answer questions must do more processing than is necessary merely to discriminate between alternatives; in generating their own answers to such questions, readers might realize whether they are retrieving information that was coded during reading or are guessing based on fragments of passage content.

To study the relation between question characteristics and awareness of comprehension, in two experiments we had university students read passages that were each followed by a single question (either short-answer or multiple-choice) about the content of the passage. The students' comprehension awareness was then assessed either by asking them to decide whether they wanted to reread the text and redo the question (Experiment 1) or by asking them to rate their certainty that their answer was correct (Experiment 2). The passages, which were selected from standardized reading comprehension tests for adults, were challenging and *inconsiderate* (Armbruster, 1984): The introductory statements were not highly informative, the passages did not contain obvious topic sentences or readily recognizable summaries, and the passages did not contain any underlining, italics, or repetition. Readers must depend more heavily on using comprehension strategies efficiently when reading inconsiderate passages than when reading well-organized, easy-to-understand passages (Pressley, 1983). Hence, inconsiderate passages seemed more appropriate for probing the metacognitive activation of rereading tactics in this study.

EXPERIMENT 1

In Experiment 1, we asked university students to read passages that were each followed by a single question (either short-answer or multiple-choice) about the content of the pas-

sage. The students then decided whether to reread the text and redo the question (if they thought there was a good chance that their initial response was in error) or move on to the next passage (if they believed their initial response was correct). The general hypothesis examined here was that answering short-answer reading comprehension questions would lead to more appropriate decisions about rereading and reprocessing of inconsiderate text than would responding to multiple-choice questions.

This hypothesis can be decomposed into two subsidiary hypotheses. First, the results of our prior study (Pressley & Ghatala, 1988) led us to expect that readers would frequently fail to look back after incorrectly answering multiple-choice questions. We reasoned that, if delusions of correctness are less likely when students respond with short answers, then the number of failures to look back after an incorrect response should be lower in the short-answer condition.

Second, we considered a corollary to the first hypothesis: Students answering multiple-choice questions might be less certain than students answering short-answer questions when they decide to move forward after correctly answering a question. After all, the familiarity of distractor content could create doubt in the minds of students who select correct answers as well as providing a false sense of security to students who select incorrect answers. Thus, we expected that regulation of rereading (i.e., moving forward after answering a question correctly and rereading after answering a question incorrectly) would be less appropriate in the multiple-choice condition than in the short-answer condition.

Method

Subjects

Thirty-four Canadian undergraduate students (23 females, 11 males); who were enrolled in a first-year psychology course served as subjects in this experiment. The students' mean age was 19.4 years (range = 18–26). Subjects were assigned randomly to either the

multiple-choice condition or the short-answer condition.

Materials

Each subject read 21 passages that were taken from practice tests for the Preliminary Scholastic Aptitude Test (PSAT: Brownstein & Weiner, 1981) and practice tests for the Scholastic Aptitude Test (SAT: Brownstein & Weiner, 1982). The experiment included 14 PSAT and 7 SAT passages, which covered literary, scientific, and social scientific topics. The readings were 118 to 520 words long. One question accompanied each passage, and either tapped a particular detail in the text (e.g., "How much of its exports does Japan ship to the United States?") or addressed the overall theme of the passage (e.g., "The message that the author wishes the reader to get is that...").

Extensive pilot testing was carried out to select questions that spanned the range from easy to difficult, and on which students' overall performance would be midrange. In addition, items were selected to ensure approximately equal difficulty between conditions. Equal difficulty was important because monitoring performance on easy tests is easier than monitoring performance on hard tests (Lichtenstein, Fischhoff, & Phillips, 1982).

Each of the multiple-choice alternatives was derived from content in the passage. Thus, subjects could not derive the correct answer to a multiple-choice question by simply discriminating between the response that referred to information in the passage and those that did not. Three of the four authors examined all multiple-choice items and concluded that the answers provided by the practice-test publisher were, in fact, discriminably more appropriate than the other choices. In the short-answer condition, all answers that addressed the question and were consistent with the meaning in the text were counted as correct. Table 1 displays a sample passage and question, the possible responses offered in the multiple-choice condition (selection (d) is correct), and sample students' responses that were counted as correct and incorrect.

After participating in the experiment, each subject completed a verbal section from the

SAT (College Entrance Examination Board, 1986). This test was composed of 10 antonym items, 10 analogy items, 5 sentence-completion items, and 15 reading comprehension questions. Thus, it provided a reliable measure of reading ability that was statistically independent of the items the subjects answered during the experiment.

Procedure

Subjects were told that they would be reading passages, one at a time. They were instructed to go through each passage carefully one time, neither rereading nor reviewing as they read. The participants were informed that their reading would be timed, but that they should not be concerned with reading quickly. The subjects were told to read at a rate that would permit them to answer a question about the passage (by completing the sentence with a short answer or choosing the answer that best completed the sentence from among multiple choices, depending on the condition). They were instructed to turn the page to the question as soon as they finished reading the passage.

Students in the short-answer condition were told to give a short answer to the sentence-completion question; students in the multiple-choice condition were told to choose the answer that best completed the sentence from the five options given. Each subject answered the question aloud, and the graduate assistant recorded the answer verbatim. The subject was then told, "If you think that your answer is probably correct, you should go to the next passage. If you think that there is a good chance that your answer is incorrect, you should go back and try to produce a better answer." Subjects who elected to look back were permitted as much time as they needed and were allowed to turn back and forth between the text and the question. No feedback about the correctness of the answer was provided at any time.

After each subject completed the 21 passages and accompanying questions, the experimenter remarked, "I noted that you looked back for (*number*) of the 21 items. How did you decide which ones to look back for?" The subject answered that question aloud and then con-

Table 1 Sample passage, question, and responses for multiple-choice and short-answer conditions: Experiment 1

Passage (both conditions)

But the weather predictions which an almanac always contains are, we believe, mostly wasted on the farmer. He can take a squint at the moon before turning in. He can "smell" snow or tell if the wind is shifting dangerously east. He can register forebodingly an extra twinge in a rheumatic shoulder. With any of these to go by, he can be reasonably sure of tomorrow's weather. He can return the almanac to the nail behind the door and put a last stick of wood in the stove. For an almanac, a zero night or a morning's drifted road—none of these has changed much since Poor Richard wrote his stuff and barns were built along the Delaware.

Question (both conditions)

The author implies that in predicting weather there is considerable value in...

Multiple-choice condition: Options presented

- (a) reading the almanac.
- (b) placing a stick of wood in the stove.
- (c) sleeping with one eye on the moon.
- (d) noting rheumatic pains.
- (e) keeping the almanac behind the door.

Short-answer condition: Sample student responses

Correct:

- Common intuition can often tell what is coming.
- way farmer feels.
- internal senses of the farmer.

Incorrect:

- knowing what each thing means in the almanac.
- farmer's interpretation of the almanac.
- value in instruments used today
- instruments are better than superstition used by people and better than the knee.

Note. Passage and multiple-choice questions are from *How to prepare for the PSAT/NMSQT* (p. 204) by S.C. Brownstein & M. Weiner, 1981, Woodbury, NY: Barron's. Copyright 1981 by Barron's. Reprinted by permission.

cluded the session by completing the verbal section of the SAT described above.

Results and discussion

The analyses are presented in the order in which subjects performed the tasks: reading the passage, answering the question, deciding whether to reread, and answering the question after rereading the passage. After these analyses, we present correlations with ability as measured by subjects' performance on the SAT verbal subtest.

Initial reading

Mean time spent by subjects in their initial reading of the texts (totaled for all 21 passages)

was 37 minutes, 36 seconds ($SD = 8$ min, 20 sec) in the short-answer condition, and 35 minutes, 7 seconds ($SD = 7$ min, 10 sec) in the multiple-choice condition. A statistical comparison between conditions (Kirk, 1982, for this and all subsequent analyses) showed no significant difference in total reading time between conditions, $t(32) = 0.93$, $p > .05$. Thus, any differences between conditions that emerge in subsequent analyses cannot be attributed to differences between conditions in reading time.

Initial performance on questions

After reading the text, subjects answered one question for each passage. Scoring these responses was straightforward in the multiple-choice condition. Two of the authors scored the short-answer data with 95 percent agreement.

The average short-answer item was an-

Table 2 Conditional probability of making appropriate decision (to reread or move on): Experiment 1

Decision	Condition		Comparison	
	Short-answer	Multiple-choice	<i>df</i>	<i>t</i>
Appropriate decision	.782	.661	20	2.79*
Forward move/correct answer	.869	.774	20	2.07
Reread/incorrect answer	.671	.525	18†	1.76

Note. Statistical analyses conducted with items as the random effect. When no significance level is indicated, $p > .05$.

* $p < .02$

†Degrees of freedom were reduced because one item was answered correctly by all short-answer subjects, and one was answered correctly by all multiple-choice subjects. As there were no incorrect confidence ratings on these items, differences between correct and incorrect ratings could not be constructed.

swered correctly, on the first try, by 56.0 percent ($SD = 28.6\%$) of the subjects, and the average multiple-choice item by 55.5 percent ($SD = 24.6\%$) of the subjects. There was no significant difference between these means, $|t(20)| = 0.06$, $p > .05$. The difficulty of items, as measured by the proportion of subjects responding correctly (e.g., Cronbach, 1949) ranged from very easy (100% correct) to very difficult (11.8% correct) in both conditions. In short, any between-condition differences in monitoring that are found cannot be due to differences in difficulty between the short-answer and multiple-choice tests. There was a significant correlation between the two conditions for item difficulty, $r = .48$, $p < .05$.

Awareness of performance

Table 2 shows the conditional probability that a subject would decide to move forward to the next passage after a correct answer, the probability that a subject would decide to reread and redo the questions after an incorrect answer, and the overall probability that a subject would make the appropriate decision (combining both of the other probabilities). Statistical comparisons between conditions were made with items as a random effect. As shown in the table, students were significantly more likely to make the appropriate decision overall in the short-answer condition than in the multiple-choice condition. There was no significant effect of condition on either of the component

probabilities, although trends favored the short-answer condition in both.

When probed later about their reasons for rereading some items, *all* subjects responded (with some variation in wording) that they looked back when they believed that they had not answered a question correctly. Many subjects emphatically stated that they had complied with this criterion, which had been given in the initial instructions.

Performance after rereading

Table 3 shows the probabilities that subjects in each condition who elected to reread and redo a question would give the correct response following their rereading. Subjects who had given the correct response after their initial reading usually gave a correct response again after rereading. Subjects who had failed to give the correct answer after the first reading produced the correct answer after rereading about half the time. As shown in the table, there was no significant difference between conditions for either of these probabilities.

Supplementary analyses by question type

Performance on questions. Post hoc inspection of the data suggested that participants responded very differently to the 15 items tapping memory for details than they did to the 6 items concerned with overall understanding of the passage. The thematic items were easier than

Table 3 Conditional probability of giving correct answer after rereading, by correctness of initial answer: Experiment 1

Correctness of first answer	Condition		Comparison	
	Short-answer	Multiple-choice	<i>df</i>	<i>t</i>
Correct	.923	.864	12†	0.86
Incorrect	.500	.565	13†	0.01

†Degrees of freedom are less than 21 because inferential statistical analyses based only on items for which there were rereadings in both conditions.

the detail items in both the short-answer condition (thematic, $M = 75.8\%$; detail, $M = 48.6\%$ correct) and the multiple-choice condition (thematic, $M = 70.7\%$; detail, $M = 49.4\%$ correct), smaller $|t(16)| = 5.23$, $p < .001$. More importantly, there was no significant difference between conditions in performance on either thematic or detail items, larger $|t(23)| = 0.78$, $p > .20$; thus, any differences in awareness between conditions for thematic or detail items are not confounded by differences in difficulty between short-answer and multiple-choice conditions.

Awareness of performance. We also calculated separate conditional probabilities that a subject would make the correct decision about rereading after answering a thematic question and after answering a detail question; these probabilities and the results of statistical comparisons between conditions are shown in Table 4. On the 15 detail items, students were again significantly more likely to make the appropriate decision in the short-answer condition than in the multiple-choice condition. On detail items, the effect of condition was also significant for the probability of rereading following an error, but not for the probability of moving forward after a correct answer. In contrast, for the 6 thematic questions, none of the analyses showed a significant effect of condition, although trends favored the short-answer condition in each. Most striking in Table 4 is the finding that subjects moved forward most of the time after attempting a thematic item, both

when they were correct and when they were incorrect.

Subjects' decision making on the 15 detail questions was compared with their decision making on the 6 thematic questions by making the six pairwise comparisons between corresponding conditional probabilities for the top and bottom halves of Table 4. The only significant effect found for question type was in appropriate rereadings following an error, which were much more likely for detail questions than for thematic questions in both the short-answer condition, $t = 3.07$, $p < .01$, and the multiple-choice condition, $t = 2.15$, $p < .05$. This effect reflects the general tendency to move forward after attempting a thematic question.

Verbal proficiency

The students in the short-answer and multiple-choice conditions were roughly equal in verbal ability: The short-answer subjects averaged 25.18 items correct ($SD = 6.05$) on the verbal section of the SAT (out of 40 possible); multiple-choice subjects averaged 21.23 items correct ($SD = 7.88$). This was not a significant difference, $t(32) = 1.64$, $p > .10$.

Performance on the verbal SAT measure varied widely, ranging from 15 to 36 in the short-answer condition and from 5 to 34 in the multiple-choice condition. Thus, meaningful correlations could be calculated involving this measure. We were particularly interested in the correlations between subjects' verbal proficiency and (a) number of passage ques-

Table 4 Conditional probability of making appropriate decision (to reread or move on) for detail and thematic questions: Experiment 1

Decision	Condition		Comparison	
	Short-answer	Multiple-choice	<i>df</i>	<i>t</i>
Detail questions				
Appropriate decision	.784	.654	14	2.65**
Forward move/correct answer	.821	.715	14	1.33
Reread/incorrect answer	.750	.598	13†	2.29*
Thematic questions				
Appropriate decision	.775	.676	5	1.01
Forward move/correct answer	.947	.875	5	2.10
Reread/incorrect answer	.269	.200	4†	0.47

Note. Statistical analyses conducted with items as the random effect. When no significance level is indicated, $p > .05$.

* $p < .05$ ** $p < .02$

†Degrees of freedom were reduced because one item was answered correctly by all short-answer subjects, and one was answered correctly by all multiple-choice subjects. As there were no incorrect confidence ratings on these items, differences between correct and incorrect ratings could not be constructed.

tions answered correctly, and (b) appropriate decision-making. There were moderate correlations between performance on the SAT verbal subtest and the number of questions answered correctly in both the short-answer condition, $r(15) = .47$, and the multiple-choice condition, $r(15) = .42$, $p < .05$ (one-tailed) for both. In contrast, the correlations between verbal proficiency and overall appropriate decision-making were not significantly greater than zero in either the short-answer condition, $r(15) = .27$, or the multiple-choice condition, $r(15) = .29$, $p > .05$ for both.

We also computed separate correlations for thematic and detail questions. For thematic questions, the correlation between verbal ability and number of passage questions answered correctly was $-.03$ (not significant, $p > .05$) in the short-answer condition and $.65$ (significant at $p < .05$) in the multiple-choice condition. For detail questions, the correlation between verbal proficiency and question performance was $.51$ ($p < .05$) in the short-answer condition and $.22$ ($p > .05$) in the multiple-choice condition. However, the $-.03$ correlation for the short-answer thematic items could not be interpreted with confidence because of very low variability

for correct recall on the thematic items ($SD = 0.79$ with 6 items). Similarly, on the multiple-choice detail questions, which correlated only $.22$ with verbal ability, 13 of the 17 subjects gave correct answers to between 7 and 11 of the items (out of 15). On the other two measures (multiple-choice thematic and short-answer detail questions), performance on the passage items was distributed normally over at least half of the scale; performance on these items did correlate significantly with verbal ability, as expected.

There was no significant correlation between verbal proficiency and comprehension awareness for thematic questions in either the short-answer condition, $r(15) = .26$, or the multiple-choice condition, $r(15) = .39$, $p > .05$ (one-tailed) for both. Nor was there any significant correlation for detail questions between verbal proficiency and comprehension awareness, $r(15) = .18$, or the multiple-choice condition, $r(15) = .12$, $p > .05$ for both.

Summary

In summary, students were significantly more likely to make the appropriate decision overall in the short-answer condition than in the

multiple-choice condition. When subjects did choose to reread, their answers improved after rereading about half the time. These effects were not dramatic, however. A more striking finding was that subjects rarely reread after incorrectly answering a thematic question, regardless of the question format. Given that participants claimed after the study that they decided to reread the passage only when they were not sure that their answers were correct, we conclude that subjects were confident about their answers to thematic questions even when those answers were in error. Inspection of the incorrect answers to the thematic questions in Experiment 1 made obvious that the errant responses were in fact related to the passages in question, although the answers were either incomplete or there was some distortion of the passage content. A reasonable hypothesis is that readers have great confidence in whatever general interpretation they have constructed while reading a passage. If so, this is a very serious deficiency in comprehension monitoring, because awareness of miscomprehension is crucial to regulation of reading and especially to regulation of text reprocessing. Thus, we examined this potential deficiency further in Experiment 2.

EXPERIMENT 2

In Experiment 2, we examined more directly the possibility that adults might not be aware that a poor answer to a thematic question is incorrect, regardless of whether the question format is short-answer or multiple-choice. We asked readers in this study to assess their confidence in their answers to thematic questions in either short-answer or multiple-choice format.

Method

Subjects

The participants in this study were 48 Canadian undergraduates (22 males, 26 females; mean age = 20.3 yrs, range = 17 to 41 yrs) recruited from the same psychology course that

supplied the participants for Experiment 1, but during the following school year. The students' mean age was 20.2 years (range = 17–41). Subjects were again randomly assigned to either the multiple-choice condition or the short-answer condition.

Materials

Each participant read 20 passages selected from the SAT verbal subtest (College Entrance Examination Board, 1988); the passages ranged from 205 to 487 words in length. As in Experiment 1, the readings covered literary, scientific, and social scientific topics. Each passage was followed by a question that tapped the overall theme of the passage. The question asked subjects to do one of the following: to state the main idea of the passage; to state its primary purpose, focus, or emphasis; to tell what the author principally wanted to discuss; or to generate a title that summarized the passage content. In the multiple-choice condition, the question was followed by five potential answers; in the short-answer condition, subjects were required to construct their own responses to the question. As in Experiment 1, pilot testing was conducted in order to identify an appropriate pool of passages and associated questions. The selected passages and accompanying questions ranged from easy to difficult, and the average question was answered correctly by about half of the participants. As in Experiment 1, we were particularly careful to select a set of items for which students' performance did not differ overall between short-answer and multiple-choice conditions.

As in Experiment 1, each multiple-choice distractor was based on elements in the passage, so that answering each question required discrimination between the best answer and other passage-related information. Two of the authors reviewed all passages and determined that the answers identified by the test publisher as correct were, in fact, the best of the multiple-choice options. In evaluating answers to the short-answer questions, the same judges accepted all answers that addressed the question and were consistent with the text. A sample passage and the accompanying question are dis-

Table 5 Sample passage, question, and responses for multiple-choice and short-answer conditions: Experiment 2**Passage (both conditions)**

Those who think of evolution as merely a process of adaptation to particular environments regard the emergence of life from the water as having been made possible by its adaptation to the dry environment outside. Though this hypothesis is not false in itself, it represents too short a view. In the long view, evolution has been toward an ever-increasing independence of the natural external environment, whatever it may be.

Let me give an example. To this day, the development of a fertilized animal cell through the embryonic stage can take place only in a liquid environment. Therefore, the first organisms to emerge on land, represented today by frogs, had to return to the water to lay their eggs. What finally brought about complete independence from the water was the development of the amniotic egg, which packaged the watery environment required by the embryo inside a membrane or shell (i.e., an amnion). So the liquid environment that a human embryo requires is provided inside an amnion inside the mother's womb. A human embryo, unlike that of a fish, is independent of the environment external to its mother, indifferent to changes in temperature—indifferent, even, to whether the environment is wet or dry.

The first human beings to emerge from the earth's atmospheric envelope, in the 1960s, were able to do so because they packaged that atmosphere in their amniotic spacesuits and spaceships. Thus, human beings are now able to visit the moon in spite of a lunar environment that would kill them on contact. What these astronauts represent is not adaptation to the natural environment but independence of it.

Question (both conditions)

The author's primary purpose in this passage was to...

Multiple-choice condition: Options presented

- (a) reveal the miracle of reproduction.
- (b) offer a wider interpretation of evolution.
- (c) explain the larger meaning of the word *amnion*.
- (d) illustrate the difference between a frog amnion and a human amnion.
- (e) compare the space age with the emergence of life from the water.

Short-answer condition: Sample student responses**Correct:**

Trying to tell us that the purpose of evolution for an animal is to become as independent as possible of the environment. Explains his version of evolution and what he thinks it is in contrast to what others believe. He believes animals become independent of their environments and not adapted to them.

Show that evolution depends on gaining independence from the environment rather than adapting to it.

Incorrect:

Talk about evolution and how everything depends on amnion. We aren't able to adapt to different environments but with amnion we are capable of surviving.

Discuss evolution and how things have adapted over time to certain circumstances.

Note. Passage and multiple-choice questions are from *10 SAT Exams* (3rd ed., p. 161) by College Entrance Examination Board, 1988. Princeton, NJ: CEEB. Copyright 1988 by CEEB. Reprinted by permission.

played in Table 5, along with the five optional responses presented in the multiple-choice condition; alternative (b) is correct. Table 5 also shows sample short-answer responses from students that were counted as correct and incorrect.

Following participation in the experimental part of the study, each subject completed a complete 40-item verbal subsection of the SAT (a different version of the SAT than the one from which the passages and questions were taken).

Procedure

The procedure of this experiment was the same as that used in Experiment 1, except that subjects in Experiment 2 were neither probed about whether they wanted to reread nor given opportunities to reread. Instead, they were asked to rate how certain they were of each answer on the following scale:

- 1 = absolutely sure answer is incorrect
- 2 = very doubtful answer is correct

- 3 = somewhat doubtful answer is correct
- 4 = 50/50 chance answer is correct
- 5 = somewhat certain answer is correct
- 6 = very certain answer is correct.
- 7 = absolutely certain answer is correct

After making a confidence rating, the subject proceeded to read the next passage.

Results and discussion

Reading time

Total time spent reading the passages averaged 36 minutes, 33 seconds ($SD = 7$ min, 11 sec) in the short-answer condition and 38 minutes, 56 seconds ($SD = 7$ min, 26 sec) in the multiple-choice condition. These times did not differ significantly, $t(46) = 1.12, p > .10$.

Performance on questions

The average short-answer item was completed correctly by 47.5 percent of the subjects ($SD = 15.6\%$); the average multiple-choice question was answered correctly by 55.2 percent of the subjects ($SD = 16.9\%$). This difference was not significant, $t(19) = 1.79, p > .05$; thus, the items in the two conditions appeared to be of equal difficulty. The easiest short-answer item was completed correctly by 79 percent of the subjects; the hardest by 21 percent. The easiest multiple-choice item was completed correctly by 75 percent of the subjects; the hardest by 21 percent. There was a significant correlation between the two conditions for item difficulty, $r(22) = .51, p < .05$.

Awareness of performance

The subjects' mean confidence ratings for their correct answers and their incorrect answers in each condition are presented in Table 6. There was no significant difference between conditions in students' confidence ratings for correct items, $|t(46)| = 1.39, p > .10$, or for incorrect items, $|t(46)| = 1.46, p > .10$. Comparison of subjects' ratings for correct items with their ratings for incorrect items showed that subjects were able to discriminate somewhat between correct and incorrect items (i.e., the difference between correct and incorrect ratings was significantly greater than zero) in both conditions, smaller $t(23) = 3.22, p < .01$. Thus, subjects could monitor somewhat whether they had answered the thematic questions correctly—to about the same degree in both conditions, $|t(46)| = 0.47, p > .05$.¹

Figures 1 and 2 compare the distribution of subjects' confidence ratings for correct items and for incorrect items in the short-answer (Figure 1) and multiple-choice (Figure 2) conditions. These figures suggest that subjects were nearly as confident in their incorrect answers as they were in their correct answers. The majority of incorrect answers in both the short-answer condition (59.7%) and the multiple-choice condition (64.4%) received confidence ratings of 5, 6, or 7 ("somewhat certain," "very certain," or "absolutely certain answer is correct").

Verbal proficiency

On the verbal section of the SAT, the mean score was 24.08 ($SD = 7.02$) for subjects in the

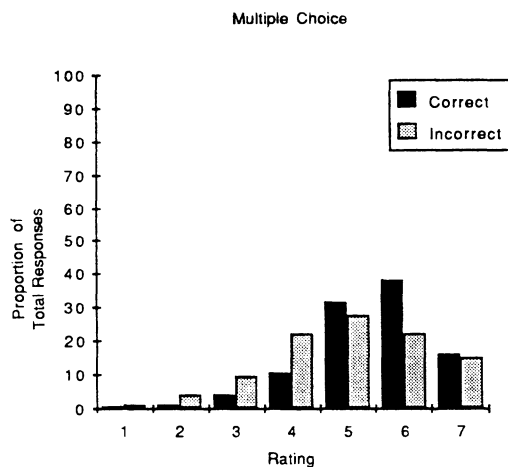
Table 6 Mean confidence rating (per subject) by correctness of answer: Experiment 2

Correctness	Short-answer		Multiple-choice	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Correct	5.20	0.61	5.43	0.53
Incorrect	4.68	0.70	4.99	0.77
Difference	0.52	0.48	0.44	0.67

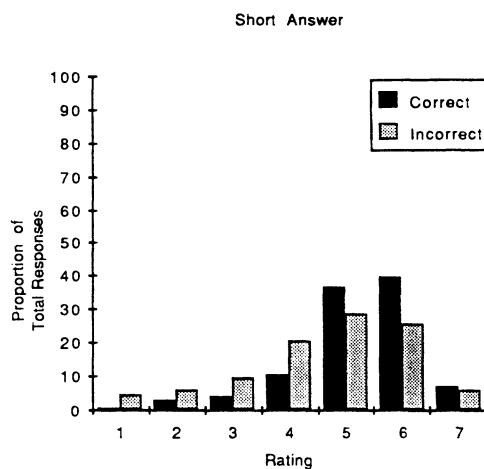
Note. Confidence rated on scale from 1 (absolutely sure answer is incorrect) to 7 (absolutely certain answer is correct).

Figure 1

Distribution of certainty ratings across subjects by correctness of response for short-answer questions: Experiment 2

**Figure 2**

Distribution of certainty ratings across subjects by correctness of response for multiple-choice questions: Experiment 2



short-answer condition, and 21.75 ($SD = 7.82$) for subjects in the multiple-choice condition, out of a maximum of 40 points. This difference between conditions was not significant, $t(46) = 1.09$, $p > .05$. As in Experiment 1, performance on this measure varied widely, ranging from 11 to 37 in the short-answer condition and from 9 to 37 in the multiple-choice condition.

This measure of verbal ability correlated significantly with objective performance on the 20 passage questions in both short-answer and multiple-choice conditions, smaller $r = .59$, $p < .01$. The correlations between verbal ability and subjects' discrimination of correct items from incorrect items were not statistically significant in either the short-answer condition, $r(22) = -.24$, or the multiple-choice condition, $r(22) = .39$, $p > .05$ for both. Thus, the pattern of significant and nonsignificant correlations in Experiment 2 was generally consistent with the pattern of significant correlations in Experiment 1. That is, although most of the SAT verbal correlations with number of passage

questions answered correctly were significantly greater than zero, all correlations between monitoring variables and verbal ability failed to reach conventional levels of significance.

Summary

In summary, students in the short-answer condition had as much difficulty discriminating correct from incorrect answers as students in the multiple-choice condition. Students in both groups had some comprehension awareness, as measured by significantly higher confidence ratings in correct than incorrect answers, although the absolute magnitudes of the differences between confidence ratings for correct and incorrect items were small. Most critically, students had high confidence in the majority of the incorrect answers in both questioning conditions. Thus, the results of Experiment 2 support the hypothesis that adults are often not aware when they are making errors in response to questions tapping the main ideas and overall themes of a passage. As expected, verbal

proficiency correlated significantly with performance measures, but not with measures of students' comprehension awareness.

GENERAL DISCUSSION

Students in both studies appeared highly confident that their incorrect answers to thematic questions were actually correct: In Experiment 1, subjects tended to move forward after answering thematic questions, even when their answers were incorrect. It seems likely that the subjects decided to move forward because they thought their answers were correct: Subjects were instructed to base their decisions on perceptions of correctness, and *all* subjects in the study reported readily that they were doing so. The high confidence ratings reported directly from subjects in Experiment 2 confirm that adults often do have a high degree of confidence that their incorrect answers to thematic questions are correct. One interpretation of the results of both experiments is that when adults are reading challenging, inconsiderate texts, they often may not be aware of gross comprehension problems—that is, they are not aware when they are not accurately comprehending main ideas.

One alternative hypothesis to explain the results of Experiment 1 is that students failed to reread after incorrectly answering a thematic question simply because continuing to read was faster and required less work than reviewing the passage to ensure that an answer was correct. Students could have realized that the fastest way to complete the task was to give an answer, claim that it was a good answer, and move on. That interpretation is at least strained by the finding that participants in Experiment 1 often did look back appropriately for detail questions. Moreover, the hypothesis that students were avoiding work is completely untenable in light of the results of Experiment 2, in which students who made low confidence ratings were not compelled to expend any additional effort; all students advanced to the next passage, regardless of their confidence rating. The most

plausible interpretation of the data of these two experiments, taken together, is that students misperceived the quality of their incorrect answers to thematic questions.

The outcomes in both studies are consistent with the high confidence in incorrect answers we reported in an earlier study (Pressley & Ghatala, 1988). In that article, we reported a distribution of confidence ratings to incorrect multiple-choice comprehension items that is strikingly similar to the distribution of ratings displayed in Figure 1. The items in that study were similar to the thematic items used here in that they could not be answered based on memory of a single detail, although—unlike the items used here—they did not require comprehension of the overall theme. However, they did require the reader to draw inferences from the text.

One of the questions posed in the present experiments was whether certain types of adjunct questions are more effective than others in prompting students to reread and restudy the text. One hypothesis based on the data reported here and in our earlier study is that questions that tap recall of factual details are more likely to prompt readers to reread than questions that require inferential comprehension. If readers are unable to retrieve a detail from memory, they appear to realize that they need to reread the text to find the necessary information. In contrast, if a question requires readers to make an inference from a passage, it may be difficult for them to evaluate their answers for adequacy: First, there is no information in the inferential question to aid evaluation of the correctness of any inferences produced. Second, subjects can often construct some type of inference based on their internal representation of the meaning of a passage, even if the internal representation is wholly inadequate to permit generation of an inference that effectively responds to the question.

As expected, based on the results of our earlier study (Pressley & Ghatala, 1988), we also found that students' comprehension awareness improved when questions were presented in a short-answer format rather than in a multiple-choice format: All short-answer versus

multiple-choice comparisons for the decision-making data in Experiment 1 favored the short-answer condition. There was also a slight trend favoring the short-answer condition in students' ability to discriminate correct from incorrect answers in Experiment 2, but the difference was not statistically significant. Thus, the short-answer format was more useful than the multiple-choice format in prompting students' comprehension awareness, but question format was a less potent factor than was the content of the question. Further studies are needed to determine precisely what features of adjunct questions are most likely to reduce (or to promote) delusions of comprehension.

The tendency toward making inappropriate decisions in Experiment 1 and toward not recognizing misperceptions in Experiment 2 was independent of the verbal ability of the respondents. The outcomes of these experiments are consistent with the results of other studies in which adults have been asked to evaluate their comprehension and/or memory of material presented in text. The general finding of studies that have used this performance prediction paradigm as an indicator of monitoring is that the correlation between verbal ability and monitoring is small. More often than not, this correlation shows that monitoring improves as reading ability increases (e.g., 3 out of 4 of the relevant correlations in Experiments 1 and 2 were in that direction); however, the relation only rarely reaches conventional levels of statistical significance (Glenberg & Epstein, 1987; Pressley & Ghatala, 1988; Pressley et al., 1987). Even those studies that show a significant relation between ability and monitoring have only found a small relation, in absolute terms (e.g., Maki & Berry, 1984).

The general failure of these studies to find that differences in monitoring by adults are related to verbal ability contrasts with a major finding in studies that have used the error detection paradigm as an indicator of monitoring. In those studies, good readers are more proficient than poor readers in recognizing when messages are flawed (e.g., August, Flavell, & Clift, 1984). Moreover, the replicable lack of ability-monitoring correlations in the performance

prediction task conflicts with the generally accepted hypothesis that metacognitive awareness increases as reading skill increases (e.g., Baker & Brown, 1984; Forrest-Pressley & Waller, 1984). The finding that there is no relation between ability to read and ability to evaluate the validity of general conclusions drawn from text has implication beyond having possibly identified a serious deficiency in monitoring. After all, if even good adult readers cannot recognize their misperceptions of a text, as seems to be the case, it is unlikely that readers can "out-grow" this deficiency with age or overcome it by improving their general reading skills. However, further studies that are focused explicitly on the nature of this deficiency are needed before firm conclusions can be drawn.

If readers of varying ability miss the point, fail to realize it, and do not reread in spite of having received as many cues to be aware of their comprehension as they did in Experiment 1, then we assume that readers are even less likely to monitor their comprehension when there are no prompts to notice comprehension problems or to reread. Thus, we conclude that when adults are reading challenging, inconsiderate texts, they often may not be aware of gross comprehension problems—that is, they are not aware that they are not accurately comprehending main ideas. Therefore, they are highly unlikely to reread or restudy the text, even when rereading would aid their comprehension. Adults might be more likely to notice whether they have missed the big ideas when they are familiar with the content area of a text, when they are processing very considerate text, or when text is fairly easy. These are all hypotheses worthy of study. However, what seems to be certain, based on the data presented here and in our prior study (Pressley and Ghatala, 1988), is that adults sometimes make gross errors in monitoring their comprehension, and that those errors certainly could undermine sophisticated regulation of reading. The challenges are to determine how great a problem this is, when such monitoring errors are likely to occur, and what can be done (if anything) to improve monitoring of main idea comprehension. The data presented here strongly support the conclusion that

simply adding adjunct questions, even if presented in the more effective short-answer format, will not improve monitoring dramatically—especially monitoring of whether overall themes are being grasped.

Implications

The opportunity to observe adults being overconfident in their erroneous interpretations of text or other communications is not confined to the psychological laboratory. For example, a major issue in the Canadian federal election in the fall of 1988 was the question of free trade with the United States. During a nationally televised debate, Prime Minister Brian Mulroney outlined his favorable opinion of the free trade treaty. Reporters from a Toronto television station asked Canadians immediately following the debate, “What did the Prime Minister say would be the effect of free trade on the lives of average Canadians?” The task of listening to Mr. Mulroney’s speech and then answering the reporter’s question is similar to the task that was presented to subjects in Experiments 1 and 2: Both groups processed a fairly inconsiderate piece of connected discourse, and then responded to a question about the general content of that discourse. The answers to the reporter’s question were diverse, even though all viewers had seen the same speech. Some of the viewers claimed that Mr. Mulroney had said that he did not care what happened to the average person, whereas others argued that the Prime Minister had specified that there would be across-the-board benefits for all Canadians.

Of course, it is not surprising that people make errors in remembering a political speech. Errors in this case probably reflect the viewers’ prior beliefs about Brian Mulroney and his positions (e.g., Neisser, 1981).² What is more important in the present context is that respondents clung tenaciously to their original answers, even when the reporter confronted them with inconsistencies between their answers and the Prime Minister’s message. The mistaken viewers really believed that their summaries were correct, just as subjects in Experiment 2 were very confident that their errant answers to the thematic questions posed in that experiment

were correct. Our hunch is that the type of overconfidence in miscomprehension captured in the studies reported here occurs in a variety of contexts. The only way to know for sure, however, is to do additional research to determine how prevalent this type of overconfidence is.

REFERENCES

- ARMBRUSTER, B.B. (1984). The problem of inconsiderate text. In G.G. Duffy, L.R. Roehler, & J. Mason, *Comprehension instruction: Perspectives and suggestions* (pp. 202-217). New York: Longman.
- AUGUST, D.L., FLAVELL, J.H., & CLIFT, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. *Reading Research Quarterly*, 20, 39-53.
- BAKER, L., & BROWN, A.L. (1984). Metacognitive skills and reading. In P.D. Pearson, M. Kamil, R. Barr, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.
- BROWNSTEIN, S.C., & WEINER, M. (1981). *How to prepare for the PSAT/NMSQT*. Woodbury, NY: Barron’s.
- BROWNSTEIN, S.S., & WEINER, M. (1982). *How to prepare for the SAT*. Woodbury, NY: Barron’s.
- COLLEGE ENTRANCE EXAMINATION BOARD (1986). *10 SAT Exams* (2nd ed.). Princeton, NJ: CEEB.
- COLLEGE ENTRANCE EXAMINATION BOARD (1988). *10 SAT Exams* (3rd ed.). Princeton, NJ: CEEB.
- EKSTRAND, B.R., WALLACE, W.P., & UNDERWOOD, B.J. (1966). A frequency theory of verbal-discrimination theory. *Psychological Review*, 73, 566-578.
- EPSTEIN, W., GLENBERG, A.M., & BRADLEY, M. (1984). Coactivation and comprehension: Contributions of text variables to the illusion of knowing. *Memory & Cognition*, 12, 355-360.
- FORREST-PRESSLEY, D.L., & WALLER, T.G. (1984). *Cognition, metacognition, and reading*. New York: Springer-Verlag.
- GARNER, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.
- GHATALA, E.S., & LEVIN, J.R. (1976). Children’s recognition memory processes. In J.R. Levin & V.L. Allen (Eds.), *Cognitive learning in children: Theories and strategies* (pp. 61-100). New York: Academic Press.
- GLENBERG, A.M., & EPSTEIN, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition*, 15, 84-93.
- GLENBERG, A.M., WILKINSON, A.C., & EPSTEIN, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, 10, 597-602.
- KIRK, R.E. (1982). *Experimental design: Procedures for the behavioral sciences*. Monterey, CA: Brooks/Cole.
- LICHTENSTEIN, S., FISCHHOFF, B., & PHILLIPS, L.D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). New York: Cambridge University Press.

- MAKI, R.H., & BERRY, S.L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 10, 663-679.
- MARKMAN, E.M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development*, 48, 986-992.
- MARKMAN, E.M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development*, 50, 643-655.
- NEISSER, U. (1981). John Dean's memory: A case study. *Cognition*, 9, 1-22.
- PRESSLEY, M. (1983). Making meaningful materials easier to learn: Lessons from cognitive strategy research. In M. Pressley & J.R. Levin (Eds.), *Cognitive strategy research: Educational applications* (pp. 239-266). New York: Springer-Verlag.
- PRESSLEY, M., & GHATALA, E.S. (1988). Delusions about performance on multiple-choice comprehension test items. *Reading Research Quarterly*, 23, 454-464.
- PRESSLEY, M., SNYDER, B.L., LEVIN, J.R., MURRAY, H.G., & GHATALA, E.S. (1987). Perceived readiness for examination performance (PREP) produced by initial reading of text and text containing adjunct questions. *Reading Research Quarterly*, 22, 219-236.
- WAGONER, S.A. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, 18, 328-346.
- WINOGRAD, P., & JOHNSTON, P. (1982). Comprehension monitoring and the error detection paradigm. *Journal of Reading Behavior*, 14, 61-76.

Footnotes

This research was supported in part by a grant to the first author from the Natural Sciences and Engineering Research Council of Canada. Correspondence regarding the article should be directed to Michael Pressley at the Department of Human Development/Institute of Child Study, College of Education, University of Maryland, College Park, MD 20742.

¹Whether a subject could discriminate correct from incorrect items was calculated here by subtracting a subject's mean rating for incorrect items from the subject's mean rating for correct items. There are alternative indices of discriminability, including per-subject gamma and point-biserial correlations (e.g., see Pressley & Ghatala, 1988). In the present case, however, there were no differences in outcome as a function of method of analysis, and thus, because the simple difference in mean ratings is easy to understand, only it is reported here.

²Some members of Canada's Progressive Conservative Party who read a preliminary version of this article expressed the belief that their leader's speech was extremely well-organized and considerate. The Liberal Party and New Democratic Party colleagues who purviewed this piece for us argued that inconsiderate was too polite a term to characterize what they viewed as logically incoherent and explicitly deceptive remarks by the Prime Minister.

Received May 20, 1988

Revision received January 24, 1989

Final revision accepted September 25, 1989

Where to submit *RRQ* articles

Through December 1990, authors should continue to submit manuscripts for the *Reading Research Quarterly* to the current editor, Dr. Philip B. Gough, *Reading Research Quarterly*, EDB 406, The University of Texas, Austin, TX 78712 (phone: 512-471-5450). Authors should continue to direct correspondence about currently pending manuscripts to Dr. Gough or his associate editors James V. Hoffman, Connie Juel, and Diane L. Schallert. Starting in January 1991, professionals wishing to submit new article manuscripts to *RRQ* should submit them to the following address, which is that of the editorial team assuming responsibility for the journal for the 1992 volume year:

Reading Research Quarterly
Ohio State University
1945 N. High Street
Columbus, OH 43210, USA.