Note Taking, Review, Memory, and Comprehension

Author(s): Mark Bohay, Daniel P. Blakely, Andrea K. Tamplin and Gabriel A. Radvansky

Source: *The American Journal of Psychology*, Vol. 124, No. 1 (Spring 2011), pp. 63-73
Published by: University of Illinois Press
Stable URL: http://www.jstor.org/stable/10.5406/amerjpsyc.124.1.0063

Note Taking, Review, Memory, and Comprehension

MARK BOHAY
University of Notre Dame

DANIEL P. BLAKELY
Florida State University

ANDREA K. TAMPLIN and GABRIEL A. RADVANSKY
University of Notre Dame

In previous work assessing memory at various levels of representation, namely the surface form, textbase, and situation model levels, participants read texts but were otherwise not actively engaged with the texts. The current study tested the influence of active engagement with the material via note taking, along with the opportunity to review such notes, and the modality of presentation (text vs. spoken). The influence of these manipulations was assessed both immediately and 1 week later. In Experiment 1 participants read a text, whereas in Experiment 2 participants watched a video recording of the material being read as a lecture. For each experiment the opportunity to take notes was manipulated within participants, and the opportunity to review these notes before the test was manipulated between participants. Note taking improved performance at the situation model level in both experiments, although there was also some suggestion of benefit for the surface form. Thus, active engagement with material, such as note taking, appears to have the greatest benefit at the deeper levels of understanding.

One of the dominant ideas about language comprehension and memory is that people represent information at multiple levels, namely the surface form, textbase, and situation model levels (van Dijk & Kintsch, 1983). In short, the surface form is a person's verbatim memory of the words and syntax used. The textbase is a representation of the abstract idea units conveyed by language apart from the surface form. Finally, the situation model is a person's referential understanding of the described events. This serves as a mental simulation and can include inferred knowledge as well as information that was explicitly presented (see also Johnson-Laird, 1983; Zwaan & Radvansky, 1998). These three levels of representation are important because they differentially contribute to memory and comprehension (Kintsch, Welsch, Schmalhofer, & Zimny, 1990; Radvansky, Zwaan, Curiel, & Copeland, 2001). The aim of the current study was to assess the influence of increased active engagement with a text via note taking, the opportunity to review such notes, and the influence of modality of presentation.

The bulk of the work on processing at these levels of representation has focused on the reading of written texts. More recent work on cognition has taken a more functionalist approach, looking at how peo-

American Journal of Psychology
Spring 2011, Vol. 124, No. 1 pp. 63–73 •

ple's interaction with the environment affects their cognition. This includes research on embodied and grounded cognition (Barsalou, 1999, 2007; Glenberg, 1997; Wilson, 2002) that shows that a person's sensorimotor interactions with the world guide thought, research on how memory has evolved to favor processing of information that is relevant for survival (Nairne & Pandeirada, 2008; Nairne, Pandeirada, & Thompson, 2008; Nairne, Thompson, & Pandeirada, 2007), and research on how the structure of the environment itself can alter basic cognitive processes, such as memory (Copeland, Magliano, & Radvansky, 2006; Radvansky & Copeland, 2006). For comprehension and memory, people can engage more actively with a text by doing more than just reading. One way to do this is by taking notes. Note taking increases the degree to which a person attends to the text, noting which ideas need to be jotted down and which are better left unnoted. Thus, by more actively engaging the reader with the material to be learned, note taking may improve memory over conditions when no notes are taken.

It is clear that note taking can improve later performance (Annis, 1975; Dyer, Riley, & Yekovich, 1979; Slotte & Lonka, 1999; Ward & Tatsukawa, 2003; see Hartley, 1983, and Hartley & Davies, 1978, for a review of this literature), both with text comprehension and spoken lecture tasks (Nye, Crooks, Powley, & Tripp, 1984; Slotte & Lonka, 1999). That said, note taking is often not an idealized representation of the information, and even successful students do not record most of the important ideas (Kiewra, 1985). So the nature of the impact it is having on later performance is unclear.

Our aim was to understand how the act of note taking, as a way to increase engagement has with information, affects processing at the various levels of representation. Several general possibilities are considered. First, it may be that the act of note taking itself creates a dual task situation that redirects cognitive resources that would otherwise be devoted to other memory and comprehension processes and sends them to the actions of scribbling something down. If so, then note taking may compromise performance, particularly at the situation model level, which requires a person to integrate information that is being presented along with inferences that are generated. Thus, note taking may orient readers more toward lower levels of processing because of increased attentional focus at this level (Peper & Mayer, 1978). This would occur at the cost of the situation model level (Zwaan, 1994).

Alternatively, it has been suggested that, during note taking, people elaborate on the material (Einstein, Morris, & Smith, 1985; Ward & Tatsukawa, 2003), with greater mental organization (Einstein et al., 1985) leading to improved memory (Craik & Lockhart, 1972). If so, then one would expect to see improved performance on the situation model level, as this is the level where such elaboration is often done.

Because notes are often intended for later review, they are, in some sense, self-generated memory cues. Therefore, a further role that note taking may play is to aid later performance (Oakhill & Davies, 1991; Slotte & Lonka, 1999), even if the notes are not one's own (Annis, 1975; Kiewra, 1985), by reminding a person of various aspects of the information. Additionally, reviewing notes can serve as another learning trial, thereby boosting memory performance. However, it is unclear at what level this assistance would occur. It could be that note review aids lower levels, such as the surface form or textbase, because this specific material is being exposed to a person once again through the act of reviewing the notes. Alternatively, it may be that review provides another opportunity for deeper processing and elaboration, which would benefit the situation model level.

Briefly, the aim of the current study was to assess the influence on memory, in terms of which level of representation is most affected, when people were more engaged during comprehension by taking notes. We also assessed at what level the notes themselves serve as memory cues to boost performance. Additionally, we assessed performance from information learned in different media, because theories of comprehension have been also been applied to video, film, and interactive video games (Copeland et al., 2006; Magliano, Radvansky, & Copeland, 2007; Zacks, Speer, & Reynolds, 2009). Finally, to address the long-term impacts of such manipulations, we tested memory both immediately after learning and 1 week later.

EXPERIMENT 1

The aim of Experiment 1 was to assess the influence of a more active engagement with a text via note taking for various levels of representation. Furthermore, because people expect to use their notes later, we also examined the impact of note review by allowing one group to review their notes. Finally, we assessed performance both immediately after reading and 1 week later.

METHOD

Participants

Ninety-seven people (53% female) were recruited from the participant pool in the Department of Psychology at the University of Notre Dame in exchange for partial course credit. Forty-eight of these people were not given the opportunity to review their notes before testing, whereas the rest were.

Materials

There were three written texts. These texts covered different topics: a summary of *Waiting for Godot* by Samuel Beckett, an overview of the biological concept of prions, and an overview of communism and Russia's economy from the 1920s to the present. We generated these texts using information from various encyclopedia and Web-based sources (they were reviewed and edited by two additional people for coherence and clarity). Although there were three texts, each participant saw only two of them. These texts were 53, 59, and 82 sentences (1,072, 1,175, and 1,392 words) long.

The memory measures were 15 four-alternative, multiple-choice recognition test questions. For each topic, there were three types of questions. First, there were questions that had verbatim answers drawn directly from the texts. Second were questions with correct answers that were paraphrases of what was stated in the text. Finally, the third group of questions had correct answers that required an inference to be made about the topic that was not explicitly stated in the text. There were five questions for each of these three types of answers. The question types were randomly ordered for each test. There were two versions of each test for each topic. One of these versions was used on the immediate test and the other for the delayed memory test. Both versions had 15 questions, with the same distribution of question types. These versions were counterbalanced across participants.

Procedure

When participants arrived, they were given consent forms and a survey for general demographic information. They were then given instructions for the experiment proper. They were told that they would be reading two texts and that for each one they would either take notes or not. In the no-notes condition, people read the text but did not take notes. In the written-notes condition, people took written notes during reading.

Participants knew that their memory would be tested. They were also told that they would be allowed to review their notes before the test. For some participants, no actual review was done, but they were told that they would have this opportunity to encourage them to take note taking seriously. The rest of the participants were given 5 min to review their notes for the one text for which notes were taken. These participants were also given 5 min to review their notes before the test after the 1-week delay.

After the notes were collected, participants took two 15-question multiple-choice tests, corresponding to the two texts that they read. The tests were presented in the same order in which the texts had been read. A follow-up session 1 week later consisted of a different set of 15 questions, which were also given in the order in which the topics were read. There was no time limit.

Analysis

To assess the influence of these manipulations on the various levels of representation, we used a variation of the Schmalhofer and Glavanov (1986) analysis (see also Fletcher & Chrysler, 1990; Kintsch et al., 1990; Radvansky, Copeland, & von Hippel, 2010; Radvansky, Copeland, & Zwaan, 2003; Radvansky et al., 2001; Zwaan, 1994). First, it should be noted that this analysis uses four types of items: verbatim items taken directly from a text; paraphrases of information that was in the text but is now worded differently, inferences of information that were not explicitly stated but are needed for the topic to be properly understood, and incorrect items that are inconsistent with a proper understanding of the material. The primary variation used in the current study was a four-alternative forced-choice test in which there was a correct answer (either a verbatim, paraphrase, or inference item) and three incorrect items.

This method uses a signal detection analysis, in which performance on different items is used to assess the different levels of representation. For the surface form, positive responses to verbatim answers are considered hits, and positive responses to paraphrases are considered false alarms. This is done not as a measure of how correct people are but as a means of assessing the degree to which people use different sources of knowledge. In this case, both the verbatim and paraphrase answers can be answered using situation model and textbase knowledge, but only the verbatim answers can be directly boosted by additional surface form knowledge because those answers were actually present in the text. Thus, by comparing performance on these two types of items we can get a measure of the added benefit of surface form memory. Similarly, a comparison of the paraphrase and inference answers provides an index of the textbase because neither of these items was actually verbatim from the text, but only the paraphrases capture ideas that were presented in the text, whereas the inferences do not. Finally, a comparison of positive responses with inferences and incorrect answers provides an index of the situation model because neither of these convey idea units that were actually in the text, but only the inferences are consistent with the information in the text and the incorrect answers are not.

RESULTS AND DISCUSSION

Prior Knowledge

The prior knowledge data showed that participants had little or no previous familiarity with the material. A scale ranging from 1 (*no knowledge*) to 5 (*extensive knowledge*) was used. The familiarity ratings were $M = 1.26$ for *Waiting for Godot,* $M = 1.30$ for prions, and $M = 1.36$ for Russia. Analyses indicated that there was no significant relationship between prior knowledge and test performance.

Memory

The memory test accuracy data (Table 1) were used to calculate the A' measures (Table 2), which were submitted to separate 2 (review or not) × 2 (delay) × 2 (note condition) mixed ANOVAs for each level of representation, with the first variable being between participants and the rest within. For the surface form measure, the only significant effect was the main effect of note condition, $F(1, 95) = 5.18$, $MSE = 0.05$, $p = .03$, with greater surface form memory when notes were taken ($M = .58$) than when they were not ($M = .52$). Essentially, when participants did not take notes, their performance was at chance. Thus, note taking does boost verbatim memory. The main effects of review and delay were not significant, $F = 1.04$ and $F < 1$, respectively, nor were any of the interactions, all Fs < 1.

For the textbase level, there were no significant main effects of review, $F(1, 95) = 2.37$, $MSE = 0.05$, $p = .13$; delay, $F < 1$; or note condition, $F < 1$. Although the two-way interactions were not significant, all ps > .10, the three-way interaction was marginally significant, $F(1, 95) = 3.38$, $MSE = 0.05$, $p = .07$. This is probably due to meaningless variation because none of these textbase scores were significantly different from chance, all ps > .30, except for the no-notes condition after a 1-week delay in the group that was allowed to review their notes. Even here, performance was significantly below chance. Thus, none of these manipulations had a unique effect on memory at the textbase level.

Finally, for the situation model level, there was a main effect of delay, $F(1, 95) = 39.10$, $MSE = 0.04$, $p < .001$, with performance worse after 1 week ($M = .62$) than when participants were tested immediately ($M = .75$). Essentially, participants had more difficulty retrieving a deeper understanding of the material a week later. There was also a main effect of having taken notes, $F(1, 95) = 7.57$, $MSE = 0.04$, $p = .007$, with performance being better when participants took notes ($M = .72$) than when they did not ($M = .66$). The interaction between these two variables was not significant, $F < 1$. Although the main effect of review was not significant, it did interact with having taken notes or not, $F(1, 95) = 4.68$, $MSE = 0.04$, $p = .03$. Essentially, as would be expected, for the topics on which participants had taken notes, being allowed to review them later ($M = .75$) served as a retrieval cue, which improved performance relative to those who did not review the notes ($M = .70$).

In summary, interacting with a text by taking notes improved performance for verbatim information and led to deeper understanding at the situation model level. There was no clear benefit at the textbase level. Note that chance performance on this measure does not mean that participants had no textbase level representation, but memory at this level did not contribute meaningfully to performance on the memory tests over and above that found at the surface form and situation model levels. In addition to the influence of note taking, the review of these notes appeared to aid performance only at the situation model level. Finally, the 1-week delay in testing affected performance only at the situation model level. This may be because the contributions of the surface form and textbase levels were so weak to begin with that they could not decline much, even after a week of forgetting.

EXPERIMENT 2

Experiment 2 was an extension of Experiment 1. However, rather than reading texts, participants received the information in a lecture format (the original texts were written to be read as lectures anyway). Furthermore, in addition to written notes, there was a third condition in which participants also took typed notes. This was done because although most research has focused on manual note taking, advances in technology have led to more students taking notes on laptop computers. Ward and Tatsukawa (2003) suggested that typing may be more effective for note taking because it allows searching, editing, sharing, and greater legibility. In addition, typing may be more efficient, less tiring, and faster than written note taking, and the faster speed of typing allows people to record more information.

METHOD

Participants

For Experiment 2, 77 people (53% female) were tested. Thirty-six were not allowed to review their notes, and the other 41 were. All participants were drawn from the participant pool in the Department of Psychology at the University of Notre Dame in exchange for partial course credit.

Materials and Procedure

The same materials were used as in Experiment 1, except that readings of the lectures were videotaped. Each lecture was 9 min long. All three lectures were presented by the same male professor. There was no additional information on the videotape (e.g., slides, blackboard writing).

When participants arrived, they were given consent forms and a survey for general demographic information. They were then given instructions for the experiment proper. They were told that they would view three 9-min videotaped lectures on the left half of a computer screen, with the right half open to a word processing program (for the computer notes condition), and that for each lecture they would take either no notes, written notes, or typed notes on a computer. The no-notes and written-notes conditions were like those in Experiment 1. In the typed-notes condition, participants took notes on the right

TABLE 1. Accuracy levels by answer type, Experiment 1

	Verbatim	Paraphrase	Inference	Wrong
		No review		
Immediate				
No notes	.70	.69	.70	.30
Written notes	.81	.71	.69	.26
Delay				
No notes	.60	.61	.57	.41
Written notes	.66	.58	.58	.39
		Review of notes		
Immediate				
No notes	.71	.66	.63	.33
Written notes	.80	.72	.74	.25
Delay				
No notes	.56	.49	.58	.45
Written notes	.69	.58	.63	.37

TABLE 2. *A'* measures (*SE*) for the various levels of representation, Experiment 1

	Surface form	Textbase	Situation model
		No review	
Immediate			
No notes	.51 (.03)	.48 (.03)	.76 (.03)*
Written notes	.58 (.03)*	.52 (.04)	.77 (.03)*
Delay			
No notes	.49 (.03)	.53 (.03)	.61 (.03)*
Written notes	.57 (.04)*	.50 (.03)	.62 (.04)*
		Review	
Immediate			
No notes	.54 (.03)	.53 (.03)	.69 (.03)*
Written notes	.57 (.03)*	.48 (.03)	.79 (.03)*
Delay			
No notes	.55 (.03)	.42 (.03)*	.58 (.04)*
Written notes	.58 (.04)*	.48 (.03)	.68 (.03)*

*$p < .05$.

half of the computer screen. All participants were in each of the three note-taking conditions, and condition and lecture type were counterbalanced across participants. Participants viewed the lectures individually on a computer. After participants viewed all the lectures, the experimenter collected the notes.

After the notes were collected, participants completed three 15-question multiple-choice tests corresponding to the three topics. The tests were presented in the same order in which the topics were viewed. As in Experiment 1, the follow-up session 1 week later consisted of a different set of multiple-choice tests, which were also given in the order in which participants viewed the topics. Participants were given an unlimited amount of time to finish these tests. Again, although all participants were told that they would be allowed to review their notes, only some actually were.

RESULTS AND DISCUSSION

Prior Knowledge

The prior knowledge data showed that the participants had little or no previous knowledge of lecture material. Again, using a scale from 1 (*no knowledge*) to 5 (*extensive knowledge*), the means were *Waiting for Godot*, $M = 1.3$; prion lecture, $M = 1.2$; and Russia lecture, $M = 1.2$.

TABLE 3. Accuracy levels by answer type, Experiment 2

	Verbatim	Paraphrase	Inference	Wrong
		No review		
Immediate				
No notes	.62	.60	.55	.41
Written notes	.72	.67	.66	.32
Typed notes	.75	.63	.64	.33
Delay				
No notes	.63	.51	.52	.45
Written notes	.58	.48	.52	.47
Typed notes	.63	.59	.61	.39
		Review		
Immediate				
No notes	.56	.64	.60	.40
Written notes	.76	.70	.64	.30
Typed notes	.67	.68	.65	.33
Delay				
No notes	.54	.57	.56	.44
Written notes	.65	.55	.61	.40
Typed notes	.67	.59	.58	.39

Memory

The memory test accuracy data (Table 3) were used to calculate the A' measures (Table 4), which were submitted to separate 2 (review or not) × 2 (delay) × 3 (note condition) mixed ANOVA, with the first variable being between participants and the rest within. For the surface form measure, there were no significant effects, although the main effect of review was marginally significant, $F(1, 75) = 3.76$, $MSE = 0.05$, $p = .06$. Surprisingly, performance was actually better overall if people were not given the opportunity to review their notes ($M = .55$) than if they were ($M = .52$). It is unclear why this was the case. Because of the oddness of this finding, and the fact that we never replicated it anywhere else, we assign little weight to it and suggest that it might be a Type I error.

There was also a marginally significant effect of note condition, $F(2, 150) = 2.37$, $MSE = 0.05$, $p = .10$. Comparisons of the different note conditions revealed marginally significant differences of the no-notes condition ($M = .51$) with the written-notes ($M = .56$) and typed-notes ($M = .55$) conditions, $F(1, 75) = 3.58$, $MSE = 0.06$, $p = .06$, and $F(1, 75) = 3.05$, $MSE = 0.04$, $p = .09$, respectively, whereas the last two conditions did not differ from one another, $F < 1$. So although there was some benefit to taking notes at the surface form level when it came to answers taken directly from the lecture, this is a weak effect. The main effect of delay was not significant, $F(1, 75) = 1.74$, $MSE = 0.05$, $p = .19$, nor were any of the interactions, all ps > .10.

For the textbase level, there were no significant main effects of review, $F < 1$; delay, $F(1, 75) = 1.98$, $MSE = 0.06$, $p = .16$; or note condition, $F < 1$. None of the interactions were significant, $F < 1$.

Finally, for the situation model level, there was a main effect of delay, $F(1, 75) = 13.79$, $MSE = 0.04$, $p < .001$, with performance growing worse over 1 week ($M = .60$) compared with immediate testing ($M = .67$). Again, people had more difficulty retrieving an understanding of the content that they had learned a week earlier. There was also a main effect of taking notes, $F(2, 150) = 6.65$, $MSE = 0.04$, $p = .002$. Comparisons of the different note conditions revealed significant differences of the no-notes condition ($M = .59$) with the written-notes ($M = .65$) and typed-notes ($M = .67$) conditions, $F(1, 75) = 6.60$, $MSE = 0.04$, $p = .01$, and $F(1, 75) = 12.17$, $MSE = 0.04$, $p = .001$, respectively, whereas the last two conditions did not differ from one another, $F = 1$. So, as in Experiment 1, there was a clear benefit to taking notes at the situation model level. Neither the main effect of review nor any of the interactions were significant, all ps > .20.

Overall the results of Experiment 2 show that note taking improved performance. Importantly, this benefit had the largest impact at the situation model level. These results support previous work that suggests that note taking can improve learning even without later review (Annis, 1975; Ward & Tatsukawa, 2003). Moreover, there was no meaningful difference between the written- and typed-notes conditions.

Analysis of Note Quantity

One concern that may arise is that people performed differently based on how extensive their notes were. To address this concern, we tallied the amount of material contained in each participant's notes and compared it with test performance in Experiments 1 and 2. This was done by first breaking the lectures down into the propositional idea units and then enumerating the number of idea units in each participant's notes (for a similar process, see Hartley & Cameron, 1967). These measures of note quantity were then compared across the various conditions and with performance on the primary task.

The mean number of idea units present in the written notes from Experiment 1 was $M = 40.4$. In Experiment 2, people in the written-notes group had a mean of $M = 36.5$ idea units, whereas those in the typed-notes group had a mean of $M = 43.6$ idea units. The mean numbers of idea units in the notes are reported in Table 5. The data for two participants in Experiment 1 were excluded because their notes were lost.

To assess whether the extensiveness of the notes was related to performance on the memory tests, the data were submitted to a correlation analysis. The relevant note quantity scores for the review and delay conditions in Experiment 1 and for the review, delay, and note conditions for Experiment 2 were compared with the surface form, textbase, and situation model level memory scores on the tests corresponding with each of these conditions. For Experiment 1, there was a significant correlation in the no-review condition with immediate situation model memory performance, $r(46) = .47$, $p = .001$. No significant correlations were found when participants were allowed to review their notes, all ps ≥ .27.

When note quantity was compared with memory test performance in Experiment 2, no significant correlations were found. However, note quantity was

TABLE 4. *A*' measures for the various levels of representation, Experiment 2

	Surface form	Textbase	Situation model
No review			
Immediate			
No notes	.52	.55	.60*
Written notes	.55	.50	.70*
Typed notes	.60*	.49	.70*
Delay			
No notes	.59*	.49	.55
Written notes	.56	.48	.54
Typed notes	.53	.47	.65*
Review			
Immediate			
No notes	.45	.53	.62*
Written notes	.55	.53	.71*
Typed notes	.49	.52	.70*
Delay			
No notes	.48	.51	.59*
Written notes	.59*	.46	.65*
Typed notes	.57	.50	.63*

*$p < .05$.

TABLE 5. Mean note quantity (in mean numbers of idea units recalled)

	Experiment 1	Experiment 2
Written notes	40.4	36.5
Typed notes	n/a	43.6
Mean	40.4	40.1

marginally correlated with immediate surface form memory, but in the negative direction, $r(41) = -.27$, $p = .086$, and immediate situation model memory in the positive direction, $r(41) = .28$, $p = .08$, when participants reviewed their notes. The remainder of the correlations did not approach significance, all $ps \geq .27$. Thus, the extensiveness of the notes showed some relation to memory performance, but the correlations were weak and inconsistent.

GENERAL DISCUSSION

This study assessed how various levels of processing are influenced by a more active interaction with the information, as with note taking. Moreover, we also looked at the impact of being able to review such notes, the retention interval, and the use of both text and video presentation on later performance.

First, this study confirmed that note taking does aid performance (see Nye et al., 1984). In both experiments, there was always evidence of some kind of benefit. This was clearly present at the situation model level, although there was also some small benefit at the surface form level. Thus, it appears that note taking facilitates deeper comprehension, consistent with previous research that suggests that people who take notes are more likely to elaborate on the material (Einstein et al., 1985; Ward & Tatsukawa, 2003).

Second, not surprisingly, the notes themselves can serve as effective memory cues. However, what is surprising is that this influence was not pervasive. It clearly appeared only at the situation model level in Experiment 1 when people read a text. With the spoken lectures in Experiment 2 it manifested itself only at the surface form level, and even then it was marginal and showed the reverse pattern. Although there may be some benefit of reviewing notes, it did not have a profound effect, at least in the context of the current study.

Third, there was also an impact of delay on performance, with people remembering less a week later than when tested immediately. However, the influence of this delay was clearly observed only at the situation model level. This may be due in part to the fact that the influence of the surface form and textbase levels on performance was generally low, in many cases at chance level, so a decline in performance after a 1-week delay would be much more difficult to observe.

Finally, the manipulation of whether the information was presented as a text or a lecture did reveal some differences on later performance, but these were minor. Looking at the overall pattern of data, it is clear that note taking had its primary influence on the situation model level, with a smaller influence at the surface form level. Thus, the patterns of data we observed here would be attributed to general comprehension processes, not to reading or listening per se.

Looking at the notes themselves, our analysis suggests that more note taking provides a later performance benefit. However, the evidence of such benefits was weak and inconsistent. Thus, at least under the circumstances used here, it seems that to improve test performance, it is of primary importance that a person be actively attending to the material by taking notes, but the amount of information noted is of secondary importance.

Overall, these results support the idea that when people are more actively engaged in material that they are trying to learn, such as by taking notes, their memory improves, particularly at deeper levels of comprehension, as with the situation model level. This is consistent with research showing that note taking can improve performance (Annis, 1975; Dyer et al., 1979; Nye et al., 1984; Slotte & Lonka, 1999; Ward & Tatsukawa, 2003). It should also be noted that if students are given handouts with parts to be filled out during comprehension lecture, this can improve memory even further (Hartley, 1976; Larson, 2009; Sambrook & Rowley, 2010), perhaps by making the material better organized and focusing students' attention on the more important concepts. More generally, the results reported here are broadly consistent with the idea that memory can be improved when people are actively engaged with the material to be learned, even if that engagement is simply the noting down of the important ideas.

More specifically, there was no support for the idea that the act of note taking produces a dual task situation that draws mental resources away from memory and comprehension processes, thereby focusing attention at the surface form and textbase levels. Instead, the fact that the bulk of the memory improvement we observed was at the situation model level is consistent with previous research suggesting that note taking may encourage people to elaborate on and better structure and organize the material being learned (Einstein et al., 1985; Ward & Tatsukawa, 2003). This is because the act of situation model creation requires a person to draw inferences about elements that are not explicitly created and to bring together various elements of the information. More generally, these findings are broadly consistent with a wide range of findings that active interaction improves memory, as with research on grounded cognition (Barsalou, 1999, 2007; Glenberg, 1997; Wilson, 2002), the benefits of survival focus encoding (Nairne & Pandeirada, 2008; Nairne et al., 2007, 2008), and event-based mental updating (Copeland et al., 2006; Radvansky & Copeland, 2006).

One concern about the current research is that all participants were repeatedly tested. Recent work has shown that prior testing can improve performance on later memory tests. This is called the testing effect (Chan, McDermott, & Roediger, 2006). The testing effect is well established, and it is present for all the conditions assessed here, and so it is a constant. If our data reflect some impact of repeated testing (thereby reducing the forgetting rate to some degree), there is no clear evidence to date that the testing effect differentially influences the various levels of representation. Moreover, the testing effect has been found to be only marginally significant ($p = .09$) when, instead of a recall test, a forced-choice recognition is used on the initial test (McDaniel, Anderson, Derbish, & Morisette, 2007), as was used here. Therefore, there is no reason to believe that the testing effect is driving the primary results reported here.

In conclusion, this study showed that performance improved when people more actively engage with the information, such as by taking notes. This improvement was observed primarily at the situation model level, where deeper understanding occurs. This increased engagement facilitated more elaborative mental processing of the material, resulting in superior situation models that allowed people to more accurately respond to subsequent memory test questions.

NOTE

Thanks to Dr. T. Merluzzi for graciously agreeing to be the lecturer and to Allie Payne for filming and editing the lectures; and to Erica Nason, Alison Dunleavy, Brad Bitterly, Josephine Kim, Katie Harris, Amanda Caravalho, Ammar Tahir, and Jessica Harrison for help with the data collection.

Address correspondence about this article to G. A. Radvansky, 218-C Haggar, Department of Psychology, University of Notre Dame, Notre Dame, IN 46556 (e-mail: gradvans@nd.edu).

REFERENCES

Annis, L. (1975). The effects of encoding and an external memory device on note-taking. *Journal of Experimental Education, 44,* 2–4.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–660.

Barsalou, L. W. (2007). Grounded cognition. *Annual Review of Psychology, 59,* 617–645.

Chan, J. C. K., McDermott, K. B., & Roediger, H. L. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135,* 553–571.

Copeland, D. E., Magliano, J. P., & Radvansky, G. A. (2006). Situation models in comprehension, memory, and augmented cognition. In M. Bernard, J. C. Forsythe, & T. Goldsmith (Eds.), *Human cognitive models in system design* (pp. 37–66). Mahwah, NJ: Erlbaum.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Dyer, J. W., Riley, J., & Yekovich, F. R. (1979). An analysis of three study skills: Note-taking, summarizing and rereading. *Journal of Educational Research, 73,* 3–7.

Einstein, G. O., Morris, J., & Smith, S. (1985). Note-taking, individual differences, and memory for lecture information. *Journal of Educational Psychology, 77,* 522–532.

Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situation models: Recognition memory for three types of textual information. *Discourse Processes, 13,* 175–190.

Glenberg, A. M. (1997). What memory is for? *Behavioral and Brain Sciences, 20,* 1–55.

Hartley, J. (1976). Lecture handouts and student notetaking. *Programmed Learning & Educational Technology, 13*(2), 58–64.

Hartley, J. (1983). Note-taking research: Re-setting the scoreboard. *Bulletin of the British Psychological Society, 36,* 13–14.

Hartley, J., & Cameron, A. (1967). Some observations of the efficiency of lecturing. *Educational Review, 20*(1), 30–37.

Hartley, J., & Davies, I. K. (1978). Note-taking: A critical review. *Programmed Learning & Educational Technology, 15*(3), 207–224.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Kiewra, K. A. (1985). Providing the instructor's notes: An effective addition to student note-taking. *Educational Psychologist, 20,* 33–39.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990).

Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29,* 133–159.
Larson, R. B. (2009). Enhancing the recall of presented material. *Computers & Education, 53,* 1278–1284.
Magliano, J., Radvansky, G. A., & Copeland, D. E. (2007). Beyond language comprehension: Situation models as a form of autobiographical memory. In F. Schmalhofer & C. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 379–392). Mahwah, NJ: Erlbaum.
McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19,* 494–513.
Nairne, J. S., & Pandeirada, N. S. (2008). Adaptive memory: Remembering with a stone-age brain. *Current Directions in Psychological Science, 17,* 239–243.
Nairne, J. S., Pandeirada, N. S., & Thompson, S. R. (2008). Adaptive memory: The comparative value of survival processing. *Psychological Science, 19,* 176–180.
Nairne, J. S., Thompson, S. R., & Pandeirada, N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 263–273.
Nye, P. A., Crooks, T. J., Powley, M., & Tripp, G. (1984). Student note-taking related to university examination performance. *Higher Education, 13,* 85–97.
Oakhill, J., & Davies, A. (1991). The effects of test expectancy on quality of note-taking and recall of text at different times of day. *British Journal of Psychology, 82,* 179–189.
Peper, R. J., & Mayer, R. E. (1978). Note taking as a generative activity. *Journal of Educational Psychology, 70,* 514–522.
Radvansky, G. A., & Copeland, D. E. (2006). Walking through doorways causes forgetting. *Memory & Cognition, 34,* 1150–1156.
Radvansky, G. A., Copeland, D. E., & von Hippel, W. (2010). Stereotype activation, inhibition, and aging. *Journal of Experimental Social Psychology, 46,* 51–60.
Radvansky, G. A., Copeland, D. E., & Zwaan, R. A. (2003). Aging and functional spatial relations in comprehension and memory. *Psychology and Aging, 18,* 161–165.
Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging, 16,* 145–160.
Sambrook, S., & Rowley, J. (2010). What's the use of webnotes? Student and staff perceptions. *Journal of Further & Higher Education, 34*(1), 119–134.
Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory and Language, 25,* 279–294.
Slotte, V., & Lonka, K. (1999). Review and process effects of spontaneous note-taking on text comprehension. *Contemporary Educational Psychology, 24,* 1–20.
van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension.* New York, NY: Academic Press.
Ward, N., & Tatsukawa, H. (2003). A tool for taking class notes. *International Journal of Human–Computer Studies, 59,* 959–981.
Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9,* 625–636.
Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General, 138,* 307–327.
Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 920–933.
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.

APPENDIX. SAMPLE TEXT AND TEST ITEMS FOR THE WAITING FOR GODOT LECTURE

Though difficult and sometimes baffling to read or (even) view, *Waiting for Godot,* Samuel Beckett's first play, was written originally in French in 1948 as *En Attendant Godot.* It premiered at a tiny theater in Paris in 1953. The play's uniqueness compelled the audiences to flock to the theaters for a spectacularly continuous four hundred performances. *Waiting for Godot* is one of the most important works of our time; revolutionizing theater in the twentieth century and having a profound influence on generations of succeeding dramatists, including such renowned contemporary playwrights as Harold Pinter and Tom Stoppard. After the appearance of *Waiting for Godot,* theater was opened to possibilities that playwrights and audiences had never before imagined. *Waiting for Godot* was a unique outburst on the literary world. It made no claim to have a place in conventional drama; rather, it carried a "fascination" of its own, authenticated by the undercurrent of resentment in accepting the illogical and unreasonable norms of the society. This play came to be considered an essential example of what Martin Esslin later called "Theater of the Absurd," a term that Beckett disavowed but which remains a handy description for one of the most important theater movements of the twentieth century. "Absurdist Theater" discards traditional plot, characters, and action to assault its audience with a disorienting experience. Characters often engage in seemingly meaningless dialogue or activities, and, as a result, the audience senses what it is like to live in a universe that doesn't "make sense." Beckett and others who adopted this style felt that this disoriented feeling was a more honest response to the post World War II world than the traditional belief in a rationally ordered universe. *Waiting for Godot* remains the most famous example of this form of drama. The play opens on a totally surreal note, with a tramp trying to pull off his boot on a lonely road under a leafless tree. There is no horizon, no sign of civilization. For a moment, this scene might even be considered comic. Eventually Vladimir enters and greets Estragon who informs Vladimir that he has spent the night in a ditch where he was beaten. They are very happy to see each other, having been separated for an unspecified amount of time. Estragon has a sore foot and is having trouble taking his boot off. . . .

Verbatim test item

This type of theater that discards traditional plot, characters, and action to assault its audience with disorienting experience is called:

a. Ridiculous
b. Absurdist (correct)
c. Bizarre
d. Strange

Paraphrase test item

________ spends the night in a ditch where he was beaten.

a. Vladimir
b. Estragon (correct)
c. Pozzo
d. Lucky

Inference test item

Audiences really responded to the play because it _________.

a. Was a unique outburst on the literary world
b. Didn't accept illogical and unreasonable norms
c. It was not a conventional drama
d. All of the above (correct)