Original Article

QJEP

Quarterly Journal of Experimental Psychology
2019, Vol. 72(12) 2807–2819

DOI: 10.1177/1747021819879434
qjep.sagepub.com
SAGE

Note-taking format and difficulty impact learning from instructor-provided lecture notes

David B Bellinger and Marci S DeCaro

Abstract

The method students use to take notes impacts how they process lecture information. The current experiment examined how the format and amount of content included in instructor-provided notes affect learning. Undergraduate students listened to a brief audio-recorded science lecture that emphasised independent facts, while using one of four note-taking guides. These guides varied in their format (outline notes, cloze notes) and level of difficulty (less-difficult, more-difficult). Outline notes included a partially complete organisational framework, promoting knowledge of relationships among concepts. Cloze notes included all lecture content with select words missing, encouraging processing of specific details. Metacognitive ratings and an objective cognitive load measure confirmed that outline note-taking was the most difficult method. However, outline notes led to higher performance than cloze notes on free recall and inference questions, and equal performance on verbatim questions. These benefits were greatest in the more-difficult outline notes condition, when less information was provided. These findings are consistent with the material-appropriate difficulty framework. Increasing note-taking difficulty was desirable, but only when the activity elicited semantic processing that complemented the type of processing afforded by the learning material.

Keywords

Learning; memory; note-taking; desirable difficulty; metacognition; material-appropriate difficulty; cognitive load

Received: 27 June 2018; revised: 15 May 2019; accepted: 8 June 2019

Lecturing is the most frequently used teaching method in higher education (Chen, 2002; Svinicki & McKeachie, 2011). Many students choose to take notes during lectures (Van Meter, Yokoi, & Pressley, 1994), even without being instructed to do so (Williams & Eggert, 2002), because they believe it will help them learn the information (Dunkel & Davy, 1989). However, students' notes are highly variable and typically inaccurate or incomplete (e.g., Kiewra, Dubois, Christensen, Kim, & Lindberg, 1989).

For these reasons, some instructors provide lecture notes, which is associated with better exam performance (e.g., Armbruster, 2009; Kiewra, 1985b). By freeing students from recording the lecture content, students can engage in more semantic processing as well as ask and answer questions during lecture (Austin, Lee, Thibeault, Carr, & Bailey, 2002). The downside of complete instructor-provided notes, however, is that students may be less likely to attend lectures (e.g., Cornelius, Owen-DeSchryver, 2008). Also, based on Fiorella and Mayer's (2016) definition of generative learning, students may receive limited benefits from complete notes because they do not encourage them to actively make sense of the information by engaging in the necessary cognitive processes (i.e., select the appropriate information, organise it by making connections between the ideas presented and the material's underlying structure, and integrate it with their prior knowledge) to bolster learning.

To counter these concerns, some instructors have adopted a modified approach: provide students with guided, but incomplete, notes. These notes typically take one of two formats: cloze notes and skeletal outlines (Austin et al., 2002; Boyle, 2012; Katayama & Robinson, 2000; Morgan, Lilley, & Boreham, 1988; Neef, McCord, & Ferreri, 2006; Williams, Weil, & Porter, 2012). *Cloze notes* include the majority of the lecture content, but essential words are

Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY, USA

Corresponding author:
Marci S DeCaro, Department of Psychological and Brain Sciences, University of Louisville, 2301 S. 3rd Street, Louisville, KY 40292, USA.
Email: marci.decaro@louisville.edu

replaced with a blank space and require students to fill in the missing words as they listen to the lecture. *Skeletal outlines* (hereafter, "outlines") provide students with an organisational framework for the lecture, requiring students to fill in the main and/or supporting ideas as they listen to the lecture. The current experiment examined how the format and completeness of instructor-provided notes impacts memory for lecture material.

Lecture note-taking (sometimes) impacts memory

Note-taking can facilitate learning at two time points: while initially taking the notes (encoding benefit) and while reviewing the notes at a later time (external storage benefit; DiVesta & Gray, 1972). The current work focused on the encoding benefits of note taking. Strategies for improving the external storage benefit of notes (e.g., spaced retrieval practice) are well documented (e.g., Karpicke & Roediger, 2010). Given the ubiquitous use of note-taking as a learning strategy, it is surprising that systematic reviews reveal only a slight encoding benefit of note-taking relative to no note-taking ($d = .26$; Kobayashi, 2005), or a lack of consensus for the encoding benefit of note-taking (Kiewra, 1985a). For example, Kiewra (1985a) found an encoding benefit in 33 out of 56 studies (59%), meaning that a sizable number of studies found no differences (21 studies; 37%) or a detrimental effect (2 studies; 4%) of note-taking. These findings suggest that any encoding benefit of note-taking may depend on important moderating factors. Drawing on the material-appropriate difficulty framework (McDaniel & Butler, 2011; McDaniel & Einstein, 1989, 2005; McDaniel, Einstein, & Lollis, 1988), we investigate two such factors: the level of difficulty induced by note-taking, and note-taking format. We examine how these factors interact to impact performance on different types of memory tests.

Understanding when difficulty during note-taking is desirable

One factor that may be important for determining the benefits of instructor-provided notes is the level of support provided, which determines how easy or difficult the note-taking task is as well as the amount of generative processing students engage in. Note-taking requires students to maintain information in their focus of attention while recording other information, which depends highly on working memory resources (Bui & Myerson, 2014; Bui, Myerson, & Hale, 2013; Engle, Tuholski, Laughlin, & Conway, 1999; Piolat, Olive, & Kellogg, 2005).

According to *cognitive load theory*, learning is enhanced if instructional activities avoid surpassing students' cognitive limits, such as providing students with a partial solution to a problem (Paas, 1992; Sweller, van Merrienboer, & Paas, 1998). Providing lecture notes to students as a learning aid may reduce the load on working memory, and thus set the conditions to allow for learning to occur. In contrast, a related concept of *desirable difficulty* emphasises the learning benefits of generating information for oneself, despite introducing greater difficulty for the learner (Bjork & Bjork, 2011; Richland, Bjork, Finley, & Linn, 2005). Considering both cognitive load theory and the desirable difficulty framework, instructors are encouraged to maximise encoding and retrieval processes that provide a productive level of difficulty without overwhelming cognitive resources (e.g., Van Merrienboer & Ayres, 2005). However, numerous studies demonstrate that not all difficulty during learning is desirable—sometimes having students generate information themselves benefits learning, and sometimes it does not (McDaniel & Butler, 2011). Furthermore, memory performance can only be explained by the underlying memory processes and neither cognitive load theory nor the desirable difficulty concept specify the memory mechanisms involved (Mitchell & Hunt, 1989). To reach this level of understanding, careful attention must be given to both the amount and type of semantic processing that students engage in while learning.

Hunt and colleagues (Einstein & Hunt, 1980; Hunt, 2003, 2013; Hunt & Einstein, 1981) provide a framework for understanding the qualitative aspect of information by distinguishing between two types of *semantic processing*: item-specific and relational processing. When students engage in *item-specific processing*, they focus on distinctive features of the information such as key terms or individual ideas. In contrast, when students engage in *relational processing*, they focus on the organisation and connections among ideas. The type of processing used determines which aspect of the information students encode in memory. Different types of learning materials, encoding tasks, and memory tests elicit or depend on item-specific or relational processing to varying extents.

McDaniel and colleagues (McDaniel & Butler, 2011; McDaniel & Einstein, 1989, 2005; McDaniel et al., 1988) proposed the *material-appropriate difficulty framework* as a guide for determining whether difficulty during learning will be beneficial. Specifically, they extend Jenkins' (1979) tetrahedral model of memory to argue that the benefits of difficulty will depend on interactions between the processing induced by four clusters of variables: the learning material, the encoding task, the memory test, and individual differences. In our study, we investigate the first three factors and control for individual differences.

According to the material-appropriate difficulty framework, memory is enhanced when students complete an encoding task that elicits semantic processing which complements (i.e., is not redundant with) the type of processing afforded by the learning material. For example, McDaniel, Einstein, Dunay, and Cobb (1986) found that students learned better from a text passage that elicited item-specific processing if the encoding task prompted relational processing, as opposed to item-specific processing. Students could better consider both the general relations between ideas (from the encoding task) as well as individuating information (e.g., specific facts or words, from the text passage), which together improves encoding and retrieval from memory (McDaniel & Einstein, 1989, 2005). Achieving material-appropriate difficulty means that both item-specific and relational processing will occur, thus supporting performance on a range of memory tests—including tests that assess item-specific details (i.e., facts directly stated in the material) as well as tests that require integration and synthesis of ideas.

Current experiment

The current experiment extended the research on material-appropriate difficulty to address how different note-taking formats and difficulty levels impact learning from a brief audio-recorded lecture. We reasoned that note-taking serves as an encoding task that can induce qualitatively different types of semantic processing. We collected several process-level measures (i.e., online cognitive load, metacognitive ratings of the note-taking experience, note completeness) to further explain the mechanisms by which note-taking format and difficulty impact learning. Although individual differences (e.g., cognitive ability) can also impact the amount of cognitive load experienced during learning, it is difficult for instructors to tailor classroom materials to account for these differences. We controlled for individual differences in prior knowledge and working memory, to examine the interaction between note-taking format and level of difficulty for students on average, beyond these individual difference factors.

Students listened to a descriptive lecture on a science topic (i.e., the components and functions of human blood). Descriptive prose (e.g., expository text or lecture) typically presents a series of independent facts with less emphasis on the underlying structure or connections among the ideas. Therefore, when learning from descriptive prose, students tend to engage in item-specific processing (i.e., treat the information as a list of independent facts; Mayer, 1985, 1987) and are less likely to process the relational aspects of the information (i.e., the organisational structure; Cook & Mayer, 1988; McDaniel et al., 1986). We used the same content from previous text-based studies which found that these descriptive materials led to item-specific processing (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011).

While listening to the lecture, students completed either cloze or outline notes. Cloze notes included an exact transcription of the lecture, but with key words missing. This task is similar to tasks used in previous studies shown to induce item-specific processing (e.g., letter-insertion task; Einstein, McDaniel, Bowers, & Stevens, 1984; McDaniel et al., 1986). In contrast, outline notes explicitly organised the prose into a hierarchy of main and supporting ideas, and required students to fill in the missing information denoted by blanks. Previous research demonstrated that using outlines induces relational processing (Einstein, McDaniel, Owen, & Cote, 1990). Because the lecture afforded item-specific processing, we therefore expected material-appropriate difficulty to be achieved when using outline notes.

We also manipulated the level of difficulty of the note-taking task. As students listened to the lecture, they were required to fill in different amounts of missing content in the cloze or outline notes. More-difficult conditions required students to complete more of the notes themselves, creating more cognitive load compared to less-difficult conditions. Students were randomly assigned to a note-taking condition in a 2 (note-taking format: Outline, Cloze) × 2 (level of difficulty: More-Difficult, Less-Difficult) design.

Memory was assessed using free recall, verbatim short answer questions, and inference short answer questions. Free recall relies on both item-specific and relational processing, and accuracy is enhanced when both types of processing occur during encoding (e.g., Einstein & Hunt, 1980; Einstein et al., 1984; Hunt & Einstein, 1981). Verbatim and inference short answer questions require item-specific and relational processing, respectively (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011). When using cloze notes, students should be engaged in only item-specific processing (from both the descriptive lecture and encoding task). In contrast, when using outline notes, students should be engaged in both item-specific (due to the descriptive lecture) and relational processing (due to the encoding task). Thus, we expected that cloze and outline notes would lead to equal performance on verbatim short answer questions, which require item-specific processing (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011). However, we expected outline notes to increase memory performance on free recall and inference questions relative to cloze notes, due to the need for relational processing on these items.

Compared to the less-difficult notes, the more-difficult notes require students to generate more information, increasing their active involvement in comprehending the lecture while also imposing greater overall cognitive load. Drawing from previous research that tested the material-appropriate difficulty framework, there is mixed evidence regarding the benefits of increasing the difficulty of the encoding task. In conditions that were set up to achieve material appropriate difficulty, increasing the difficulty of the encoding task benefitted learning of word lists (McDaniel et al., 1988) but hindered memory performance when learning from a text passage (Einstein et al., 1990). One interesting note about these studies is that the learning material itself was manipulated by presenting it in a random order or withholding some information from the student. In our experiment, however, the lecture was intact and only the learning aid was manipulated to induce different types of semantic processing. We assumed that this additive approach (i.e., note-taking is supplementary to the lecture) could accommodate additional difficulty and thus we predicted that increasing difficulty will benefit learning—but only in the outline notes condition.

In the More-Difficult Outline condition, students engage in both item-specific and relational processing, while also generating more information for themselves. In contrast, by decreasing difficulty, Less-Difficult Outline notes may decrease active processing, reducing learning compared to the More-Difficult Outline condition (e.g., Morgan et al., 1988). Importantly, both the degree and type of processing should matter—despite increasing encoding difficulty, the More-Difficult Cloze notes should lead to only item-specific processing. Thus, both Cloze notes conditions should lead to lower learning than the More-Difficult Outline notes condition. This prediction is consistent with findings when learning from word lists as well as text passages (McDaniel & Butler, 2011; McDaniel & Einstein, 1989, 2005; McDaniel et al., 1988).

By examining the combined impact of note-taking format and level of difficulty, this study may reveal important factors moderating the benefits of instructor-provided notes. In addition, this study applies the material-appropriate difficulty framework to a common learning situation (i.e., lecture note-taking) by providing the first empirical test to examine the interactions between different types of semantic processing (i.e., induced by the encoding task and learning material) and levels of difficulty in an encoding task that is not self-paced.

Method

Participants

Participants were undergraduate students enrolled in psychology courses ($N=123$, M age $=20.24$ years, $SD=3.36$, 63.4% female). The majority of students identified themselves as White (81%), with the remaining individuals identifying themselves as Black (10%), Asian (5%), Hispanic or Latino (1%), or other (4%). Additional students were tested, but excluded from analyses, for three reasons. First, students were excluded for not following experiment instructions: (a) completing less than 30% of the notes handout, indicating that they were not sufficiently exposed to the processing manipulation of the note-taking format ($n=2$), (b) committing 20% or more errors on the automated reading span task ($n=4$; Conway et al., 2005), or (c) missing more than two responses across both reaction time tasks ($n=7$). Second, students were excluded for having a high level of prior knowledge, and therefore less to learn from the lecture, by producing at least 50% of the components or functions of human blood on the prior knowledge question ($n=47$). Third, students were excluded for reporting that English was not their first language, and thus may not fluently understand the verbal or written materials ($n=5$).

Experimental design

Students were randomly assigned to a note-taking condition in a 2 (note format: Cloze, Outline) × 2 (level of difficulty: Less-Difficult, More-Difficult) between-subjects factorial design (Less-Difficult Cloze $n=28$, More-Difficult Cloze $n=33$, Less-Difficult Outline $n=32$, More-Difficult Outline $n=30$).

Materials

Prior knowledge assessment. Students completed a brief cued-recall question assessing their prior knowledge. Specifically, students were asked to list each component of human blood and its corresponding function. Students were given the option to write "I don't know." Below the prompt were eight lines, labelled, "Component 1, Function 1, etc." for each of four component/function pairs. One point was given for each correctly listed component (i.e., plasma, red blood cells/erythrocyte, white blood cells/leukocyte, platelets) and function (eight points total).

Lecture. Students listened to a 2-min science lecture about the components and functions of human blood (adapted from text-based studies), which affords item-specific processing (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011). The 245-word lecture was presented at an average rate of 117 words per minute and included 33 individual idea units (i.e., a small group of words that represent a single idea or fact). Idea units were used to assess note-taking and free recall performance.

Online cognitive load. Students were asked to wear headphones and complete a reaction time measure of online cognitive load (Piolat et al., 2005) by pressing the spacebar as quickly as possible, using their non-writing hand, once they heard a tone. The 2-min task presented six auditory tones at predetermined random intervals ranging from 15 to 30 s. Students first completed this task as a single-task baseline measure, then as a secondary task during the lecture. The tones never overlapped with words in the lecture.

Reaction times were calculated by subtracting the onset time for each tone from the time at which the space bar was pressed. If no response time to a tone was recorded, then the missing response time was replaced with the maximum time allotted to respond to the corresponding tone. Then, the onset time for the tone was subtracted from the replaced response time to calculate the reaction time.

Median interference in reaction time (IRT) scores were calculated by (a) subtracting the baseline reaction time from the secondary-task reaction time and then (b) calculating the median value of the six reaction time differences. A positive median IRT indicates an increase in reaction time (i.e., slower response) during the secondary task relative to the baseline task and may be interpreted as an increase in cognitive load induced by lecture note-taking (Piolat et al., 2005).

Note-taking. Figure 1 illustrates each type of note-taking handout (cloze, outline) for each level of difficulty (less, more). The handouts were designed to simulate two types of instructor-provided partial notes. *Cloze notes* handouts provide a transcription of the lecture but with key word(s) deleted from each idea unit. Students were asked to fill in the missing words as they listened to the lecture (e.g., Einstein et al., 1984; McDaniel et al., 1986). *Outline notes* handouts explicitly organised the prose into a hierarchy of main and supporting ideas. Students were asked to fill in the missing information denoted by blanks (Einstein et al., 1990).

The difficulty of the note-taking task was manipulated by requiring students to fill in different amounts of missing content. In the Less-Difficult Cloze and Less-Difficult Outline handouts, approximately half (i.e., 17) of the idea units were missing information (the same 17 words were missing in both handouts). In the More-Difficult Cloze and More-Difficult Outline handouts, all 33 idea units were incomplete.

Working memory capacity. The Automated Reading Span Task (Redick et al., 2012) served as a distractor task between the lecture and memory tests. Scores were also used as a covariate to estimate effects of condition independent of individual differences in working memory capacity. Students were asked to judge whether a sentence made sense and then presented with a letter for recall at the end of the set. After three to seven sentence-letter combinations, students were presented with a recall grid and asked to select the letters they saw during the trial in the correct order. Students viewed three administrations for each set size (i.e., 75 total sentence-storage pairs). The total score was calculated by summing the total number of correct responses out of 75 (Conway et al., 2005).

Questionnaires. The *post-lecture questionnaire* included four questions assessing students' metacognition. Using a 7-point Likert-type scale, we assessed students' perceptions of how (a) difficult (1 = *very easy*, 7 = *very difficult*; adapted from DeLeeuw & Mayer, 2008), (b) helpful (1 = *not at all helpful*, 7 = *very helpful*), and (c) enjoyable (1 = *not at all enjoyable*, 7 = *very enjoyable*) the note-taking task was, and (d) their perceived comprehension of the lecture (1 = *did not comprehend it very well*, 7 = *comprehended it very well*; adapted from Einstein et al., 1990).

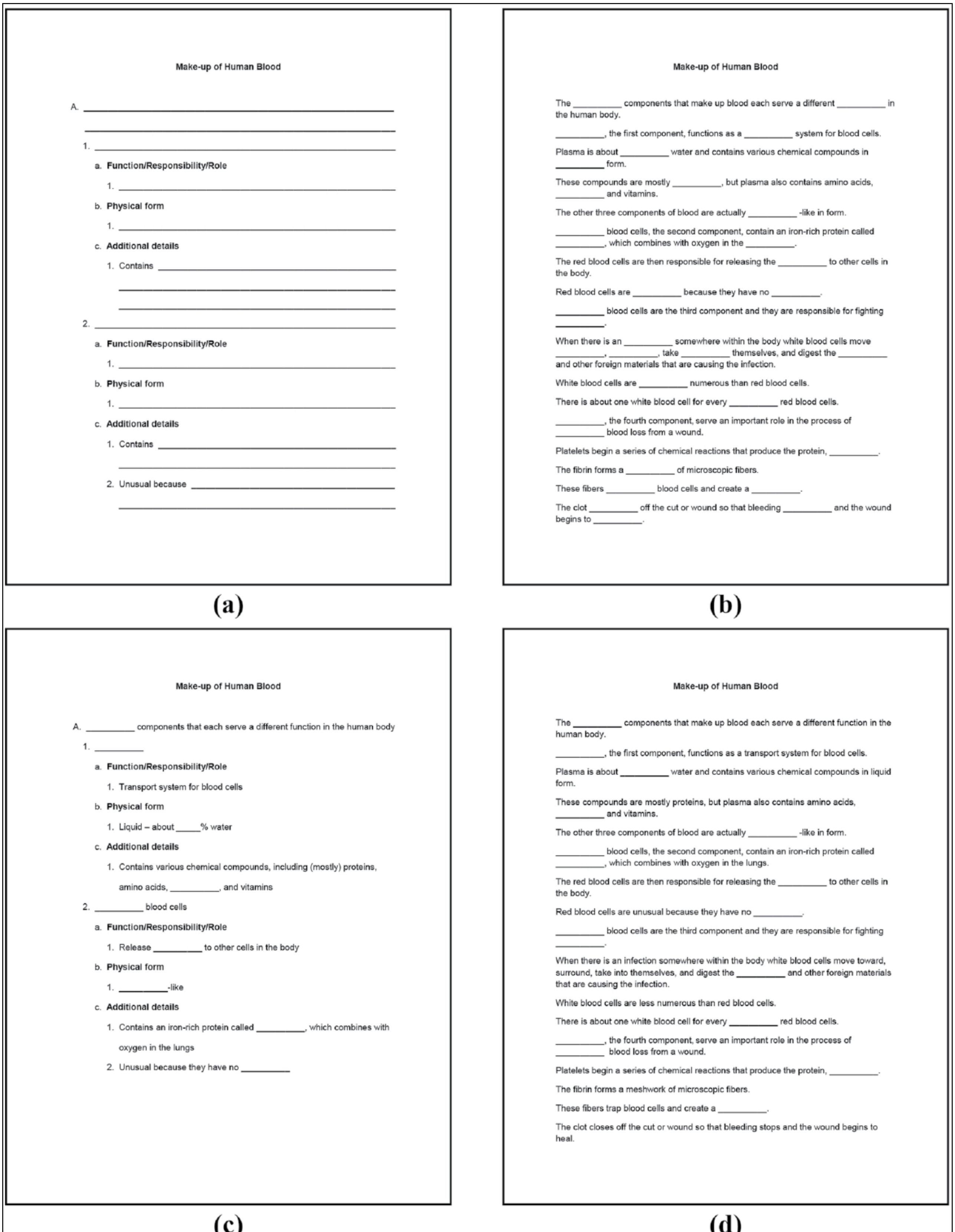
The *post-experiment questionnaire* requested demographic information and students' note-taking preferences (i.e., a 9-option multiple choice question asking which approach to note-taking they use most often during lectures in science courses) and experience with instructor-provided notes during their post-secondary education (i.e., two 7-option multiple choice questions asking for how many courses their college instructors have provided them with partial and complete lecture notes and one 3-option multiple choice question asking when they could first access the instructor-provided notes).

Memory tests. To assess learning from the lecture, two types of memory tests were administered on a computer. The *free recall test* (Karpicke & Blunt, 2011) required students to document everything they could remember from the lecture, within a 7 min time limit. The *short answer test* (adapted from Blunt & Karpicke, 2014; Karpicke & Blunt, 2011) included 10 verbatim and 4 inference questions. The *verbatim questions* (Cronbach's $\alpha = .52$) asked for specific facts (typically a single idea unit) stated directly in the lecture and thus assessed recall of item-specific information (Blunt & Karpicke, 2014). For example, the question "What percentage of plasma is water?" corresponded to the idea unit "Plasma is about 90% water." In contrast, the *inference questions* (Cronbach's $\alpha = .48$) assessed recall of relational information, because they required students to connect multiple idea units from the lecture to reason beyond the information provided (Blunt & Karpicke, 2014). For example, the question "What would happen to the blood flow from a wound if the body had no fibrin?" referred to the following idea units: (a) "The fibrin forms a meshwork of microscopic fibres"; (b) "These fibres trap blood cells"; (c) "and create a clot"; (d) "The clot closes off the cut or wound"; (e) "so that bleeding stops."

Students were required to spend a minimum of 15 s answering each short answer question (adapted from Karpicke & Blunt, 2011). No maximum time was given. Two raters scored 20% of all memory tests and notes, demonstrating high consistency between raters (Cohen's kappas = .91 to 1.00). Remaining memory tests and notes were scored by only one rater.

Procedure

Students completed the study individually in separate rooms. After providing informed consent, students were asked to complete the prior knowledge assessment. Then, students were informed that they would listen to a lecture about the human body, and that they would not be able to rewind, fast-forward, or pause the recording. They were instructed to take notes using the handout provided to them in order to help them learn the information for a memory test at the end of the experiment. Students were then asked to wear headphones and complete the baseline reaction time task.



(a)

Make-up of Human Blood

A. __________

1. __________
   a. **Function/Responsibility/Role**
      1. __________
   b. **Physical form**
      1. __________
   c. **Additional details**
      1. Contains __________

2. __________
   a. **Function/Responsibility/Role**
      1. __________
   b. **Physical form**
      1. __________
   c. **Additional details**
      1. Contains __________
      2. Unusual because __________

(b)

Make-up of Human Blood

The _________ components that make up blood each serve a different _________ in the human body.

_________, the first component, functions as a _________ system for blood cells.

Plasma is about _________ water and contains various chemical compounds in _________ form.

These compounds are mostly _________, but plasma also contains amino acids, _________ and vitamins.

The other three components of blood are actually _________ -like in form.

_________ blood cells, the second component, contain an iron-rich protein called _________, which combines with oxygen in the _________.

The red blood cells are then responsible for releasing the _________ to other cells in the body.

Red blood cells are _________ because they have no _________.

_________ blood cells are the third component and they are responsible for fighting _________.

When there is an _________ somewhere within the body white blood cells move _________, _________, take _________ themselves, and digest the _________ and other foreign materials that are causing the infection.

White blood cells are _________ numerous than red blood cells.

There is about one white blood cell for every _________ red blood cells.

_________, the fourth component, serve an important role in the process of _________ blood loss from a wound.

Platelets begin a series of chemical reactions that produce the protein, _________.

The fibrin forms a _________ of microscopic fibers.

These fibers _________ blood cells and create a _________.

The clot _________ off the cut or wound so that bleeding _________ and the wound begins to _________.

(c)

Make-up of Human Blood

A. _________ components that each serve a different function in the human body

1. _________
   a. **Function/Responsibility/Role**
      1. Transport system for blood cells
   b. **Physical form**
      1. Liquid – about _____% water
   c. **Additional details**
      1. Contains various chemical compounds, including (mostly) proteins, amino acids, _________, and vitamins

2. _________ blood cells
   a. **Function/Responsibility/Role**
      1. Release _________ to other cells in the body
   b. **Physical form**
      1. _________-like
   c. **Additional details**
      1. Contains an iron-rich protein called _________, which combines with oxygen in the lungs
      2. Unusual because they have no _________

(d)

Make-up of Human Blood

The _________ components that make up blood each serve a different function in the human body.

_________, the first component, functions as a transport system for blood cells.

Plasma is about _________ water and contains various chemical compounds in liquid form.

These compounds are mostly proteins, but plasma also contains amino acids, _________ and vitamins.

The other three components of blood are actually _________ -like in form.

_________ blood cells, the second component, contain an iron-rich protein called _________, which combines with oxygen in the lungs.

The red blood cells are then responsible for releasing the _________ to other cells in the body.

Red blood cells are unusual because they have no _________.

_________ blood cells are the third component and they are responsible for fighting _________.

When there is an infection somewhere within the body white blood cells move toward, surround, take into themselves, and digest the _________ and other foreign materials that are causing the infection.

White blood cells are less numerous than red blood cells.

There is about one white blood cell for every _________ red blood cells.

_________, the fourth component, serve an important role in the process of _________ blood loss from a wound.

Platelets begin a series of chemical reactions that produce the protein, _________.

The fibrin forms a meshwork of microscopic fibers.

These fibers trap blood cells and create a _________.

The clot closes off the cut or wound so that bleeding stops and the wound begins to heal.

Figure 1. Examples of note-taking handouts: (a) More Difficult Outline notes, (b) More Difficult Cloze notes, (c) Less Difficult Outline notes, and (d) Less Difficult Cloze notes.

Table 1. Students' self-reported experiences with instructor-provided notes.

	Complete notes	Partial notes
0 courses	10.6%	47.2%
1–2 courses	22.8%	33.3%
3–4 courses	31.7%	13.0%
5–6 courses	9.8%	4.1%
7–8 courses	7.3%	1.6%
9–10 courses	3.3%	–
11 + courses	14.6%	0.8%

Immediately before the lecture began, students received an example of the type of handout they would use to take notes, to familiarise them with the format. Students were shown both a blank and completed example handout, on a different topic (i.e., the human ear). Students were directed to focus on the format of the notes rather than the content, and were provided with 60 s to review the example handout before they were given the handout to be used during the lecture.

Students then performed the reaction time task during the lecture while taking notes, as a measure of online cognitive load. After the lecture, the experimenter collected the notes, and students completed the post-lecture questionnaire, followed by the working memory capacity task. Next, students completed the post-experiment questionnaire, followed by the memory tests for the lecture. Finally, students were debriefed. The experiment lasted approximately 50 min.

Results

Note-taking habits and experiences

In order to gauge experience with note-taking, we first examined students' self-reported note-taking in science courses. Nearly all students (97.6%) reported taking notes during lectures. The majority reported creating outlines (26.0%) or writing a list of bullet points (39.8%). The remaining students reported that they try to write down everything the instructor says (21.1%), use a copy of the instructor's PowerPoint slides to guide their note-taking (4.9%), or responded that they use a method other than any of the above (5.7%). These findings suggest that students are generally familiar with outlining and value having notes that are as complete as possible. Additionally, no students reported using the Cornell note-taking method (e.g., Quintus, Borr, Duffield, Napoleon, & Welch, 2012), graphic organisers (e.g., Ponce & Mayer, 2014), or matrix notes (e.g., Kiewra, Benton, Kim, Risch, & Christensen, 1995).

As shown in Table 1, students indicated that professors were more likely to provide a complete copy of the lecture information (e.g., PowerPoint slides) compared to a partially complete copy of the lecture information (e.g., key terms or definitions deleted). Thus, in this sample, students were less likely to have had experience with partially completed notes. Students reported that it is more common for them to first access instructor-provided notes before lecture ($n = 113$; 463 courses indicated) relative to after lecture ($n = 63$; 207 courses indicated).

Preliminary analyses

Separate 2 (note format: cloze, outline) × 2 (level of difficulty: less-difficult, more-difficult) between-subjects factorial analyses of variance (ANOVAs) revealed that neither working memory capacity nor prior knowledge differed based on note format, working memory capacity, $F < 1$; prior knowledge, $F(1, 119) = 1.61$, $p = .207$, $\eta_p^2 = .01$; or level of difficulty, working memory capacity, $F(1, 119) = 2.99$, $p = .086$, $\eta_p^2 = .03$, prior knowledge, $F(1, 119) = 2.41$, $p = .123$, $\eta_p^2 = .02$. No interactions were found, $F\text{s} < 1$, indicating that the four note-taking conditions were statistically equivalent in working memory capacity and prior knowledge.

Subsequent analyses controlled for working memory capacity ($M = 57.12$, $SD = 9.42$) and prior knowledge ($M = 0.12$, $SD = 0.15$), to demonstrate condition effects while accounting for these individual differences.[1] Table 2 presents the main effects of working memory capacity and prior knowledge for each model reported below.

Primary analyses

Separate 2 (note format: cloze, outline) × 2 (level of difficulty: less-difficult, more-difficult) between-subjects factorial ANCOVAs were used to analyse memory performance, cognitive load, note completeness, and metacognitive ratings for the lecture. The descriptive nature of the lecture is thought to encourage item-specific processing (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011). Thus, we hypothesised that taking outline notes, which is thought to encourage relational processing (Einstein et al., 1990), would facilitate material-appropriate processing and benefit memory performance on free recall and short answer inference questions relative to cloze notes. Moreover, based on the material-appropriate difficulty account (McDaniel & Butler, 2011), we predicted an interaction, with the highest performance on these items in the More-Difficult Outline condition. We hypothesised that there would be no difference between conditions on short answer verbatim questions.

Memory tests

Free recall. Level of difficulty did not significantly impact free recall, $F(1, 117) = 1.65$, $p = .201$, $\eta_p^2 = .01$, $d = .11$. As predicted, a main effect of note format indicated that outline notes ($M = 0.25$, $SE = 0.01$) led to superior free

Table 2. Main effects of working memory capacity and prior knowledge for each analysis.

	Working memory capacity		Prior knowledge	
	F	η_p^2	*F*	η_p^2
Free recall	2.01	.02	16.35***	.12
Verbatim (short answer)	1.99	.02	11.07**	.09
Inference (short answer)	0.07	.00	10.00**	.08
Online cognitive load	0.13	.00	3.02	.03
Note completeness	1.45	.01	3.77	.03
Difficult	4.27*	.04	0.39	.00
Enjoyable	0.51	.00	0.17	.00
Comprehend	0.06	.00	8.11**	.07
Helpful	3.42	.03	0.44	.00

*$p<.05$; **$p<.01$; ***$p<.001$.

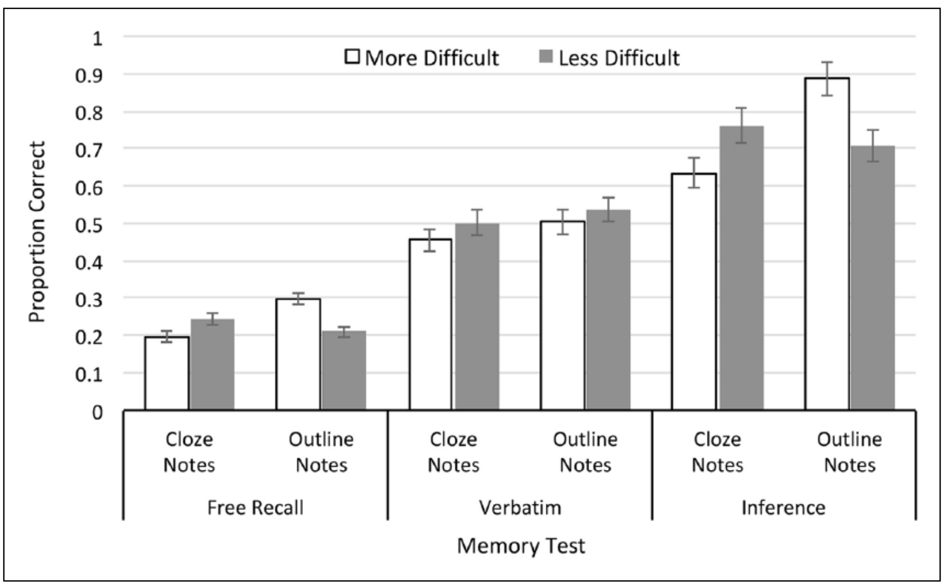


Figure 2. Mean proportion correct on free recall and short answer tests (verbatim, inference) as a function of note format and level of difficulty. Error bars represent ± 1 standard error of the mean.

recall compared to cloze notes ($M=0.22$, $SE=0.01$), $F(1, 117)=4.77, p=.031, \eta_p^2=.04, d=.32$. This effect was qualified by a significant interaction, $F(1, 117)=19.95$, $p<.001$, $\eta_p^2=.15$. As shown in Figure 2 and Table 3, students who completed More-Difficult Outline notes produced more total idea units during free recall compared to students who completed More-Difficult Cloze, $F(1, 59)=20.44, p<.001, \eta_p^2=.26, d=1.10$; Less-Difficult Cloze, $F(1, 54)=6.06, p=.017, \eta_p^2=.10, d=.50$; and Less-Difficult Outline notes, $F(1, 58)=17.68, p<.001, \eta_p^2=.23$, $d=.94$. These findings suggest that completing more-difficult outline notes promotes better learning than the other note-taking conditions.

Short answer questions. As predicted, on verbatim questions, there were no main effects of note format, $F(1, 117)=1.59$, $p=.210, \eta_p^2=.01, d=.21$; level of difficulty, $F(1, 117)=1.52$, $p=.220, \eta_p^2=.01, d=.32$; or an interaction, $F<1$ (Figure 2). On inference questions, there was no main effect of level of difficulty, $F<1$, but there was a main effect of note format, $F(1, 117)=5.14$, $p=.025$, $\eta_p^2=.04$, $d=.35$. Outline notes ($M=0.80$, $SE=0.03$) led to more accurate inferences compared to cloze notes ($M=0.70$, $SE=0.03$). This effect was qualified by a significant interaction, $F(1, 117)=12.31$, $p=.001$, $\eta_p^2=.10$. As shown in Figure 2 and Table 3, completing More-Difficult Outline notes led to more accurate inferences compared to all of the other guided notes, More-Difficult Cloze, $F(1, 59) =18.85$, $p<.001$, $\eta_p^2=.24$, $d=1.10$; Less-Difficult Cloze, $F(1, 54) =5.03$, $p=.029$, $\eta_p^2=.09$, $d=.45$; Less-Difficult Outline, $F(1, 58) =6.92$, $p=.011$, $\eta_p^2=.11$, $d=.62$.

Online cognitive load

Completing outline notes ($M=240.19$, $SE=16.45$) increased secondary task reaction time compared to cloze notes

Table 3. Mean memory test scores (proportion correct), metacognitive ratings (out of 7), note completeness (proportion correct), and online cognitive load (median interference in reaction time—ms) as a function of note format and level of difficulty. Standard deviations are in parentheses.

	Cloze notes		Outline notes		Note format		Level of difficulty	
	Less difficult (n=28)	More difficult (n=33)	Less difficult (n=32)	More difficult (n=30)	Cloze notes (n=61)	Outline notes (n=62)	Less difficult (n=60)	More difficult (n=63)
Memory tests								
Free recall	0.25 (0.08)	0.19 (0.10)	0.21 (0.09)	0.29 (0.08)	0.22 (0.10)	0.25 (0.09)	0.23 (0.09)	0.24 (0.10)
Verbatim	0.52 (0.19)	0.45 (0.18)	0.54 (0.19)	0.49 (0.19)	0.48 (0.19)	0.52 (0.19)	0.53 (0.19)	0.47 (0.18)
Inference	0.78 (0.23)	0.64 (0.24)	0.71 (0.32)	0.87 (0.17)	0.70 (0.25)	0.79 (0.27)	0.74 (0.28)	0.75 (0.24)
Metacognitive ratings								
Difficult	1.71 (1.18)	2.49 (1.23)	3.56 (1.32)	4.73 (1.26)	2.13 (1.26)	4.13 (1.41)	2.70 (1.55)	3.56 (1.67)
Enjoyable	4.68 (1.19)	4.30 (1.40)	4.84 (1.51)	3.50 (1.41)	4.48 (1.31)	4.19 (1.60)	4.77 (1.36)	3.92 (1.45)
Comprehend	5.29 (1.54)	4.58 (1.48)	4.84 (1.48)	4.13 (1.43)	4.90 (1.54)	4.50 (1.49)	5.05 (1.51)	4.37 (1.46)
Helpful	4.36 (1.68)	3.27 (1.46)	4.34 (1.23)	3.63 (1.43)	3.77 (1.65)	4.00 (1.37)	4.35 (1.45)	3.44 (1.45)
Process measures								
Note completeness	1.00 (0.02)	0.96 (0.04)	0.81 (0.13)	0.52 (0.12)	0.97 (0.04)	0.67 (0.19)	0.90 (0.13)	0.75 (0.23)
Online cognitive load	88.39 (109.41)	159.61 (125.49)	215.84 (143.94)	269.82 (134.76)	126.92 (122.74)	241.96 (141.08)	156.37 (143.09)	212.09 (140.36)

Values in the table are raw scores, and do not account for covariates used in the primary analyses.

(M=126.71, SE=16.63), F(1, 117)=23.31, $p<.001$, $\eta_p^2=.17$, d=.87, indicating greater cognitive load in the outline notes conditions. In addition, notes with more difficulty (M=212.53, SE=16.44) increased reaction time relative to notes with less difficulty (M=154.37, SE=16.88), F(1, 117)=5.96, p=.016, $\eta_p^2=.05$, d=.39, indicating increased cognitive load in the More-Difficult conditions. No interaction was found, $F<1$.

Note completeness

Note completeness was measured as the proportion of correct words written in the correct blanks. Cloze notes (M=0.97, SE=0.01) were more complete than outline notes (M=0.67, SE=0.01), F(1, 117)=355.53, $p<.001$, $\eta_p^2=.75$, d=2.18. Less-difficult notes (M=0.90, SE=0.01) were more complete than more-difficult notes (M=0.74, SE=0.01), F(1, 117)=97.95, $p<.001$, $\eta_p^2=.46$, d=.80. These effects were qualified by an interaction, F(1, 117)=58.93, $p<.001$, $\eta_p^2=.34$. Even though More-Difficult Outline notes led to higher memory performance, these notes (M=0.53, SE=0.02) were less complete than in each other condition, More-Difficult Cloze (M=0.95, SE=0.02), F(1, 59)=369.43, $p<.001$, $\eta_p^2=.86$, d=5.02; Less-Difficult Cloze (M=0.99, SE=0.02), F(1, 54)=397.71, $p<.001$, $\eta_p^2=.88$, d=5.44; Less-Difficult Outline (M=0.81, SE=0.02), F(1, 58) =81.75, $p<.001$, $\eta_p^2=.59$, d=2.32. Overall, there was a negative partial correlation between note completeness and free recall, r_p(119)=-.21, p=.019, and inference memory scores, r_p(119)=-.19, p=.038. No relationship was found between note completeness and verbatim scores, r_p(119)=.10, p=.272. These findings mirror the overall results, in that outline notes were less complete and also led to higher free recall and inference scores.

Metacognitive ratings

Difficult. Outline notes (M=4.14, SE=0.16) were rated as more difficult than cloze notes (M=2.11, SE=0.16), F(1, 117)=81.00, $p<.001$, $\eta_p^2=.41$, d=1.50. In addition, notes taken in the More-Difficult conditions (M=3.65, SE=0.16) were rated as more difficult than those taken in the Less-Difficult conditions (M=2.59, SE=0.16), F(1, 117)=21.67, $p<.001$, $\eta_p^2=.16$, d=.53. No interaction was found ($F<1$; Figure 3; Table 3).

Enjoyable. Students enjoyed using the Less-Difficult notes (M=4.74, SE=0.18) compared to the More-Difficult notes (M=3.92, SE=0.18), F(1, 117)=10.02, p=.002, $\eta_p^2=.08$, d=.60. There was no main effect of note format, F(1, 117)=1.59, p=.210, $\eta_p^2=.01$, d=.20, or interaction, F(1, 117)=3.75, p=.055, $\eta_p^2=.03$ (Figure 3; Table 3).

Comprehend. Students who completed Less-Difficult notes (M=5.01, SE=0.19) thought they comprehended the

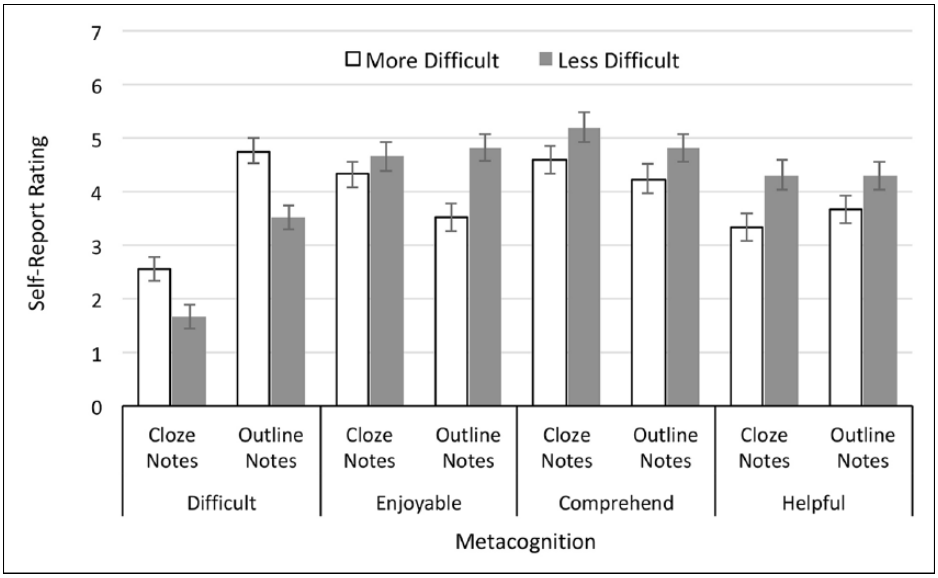


Figure 3. Mean self-report ratings of metacognitive factors as a function of note format and level of difficulty. Error bars represent ± 1 standard error of the mean.

lecture better than students who completed More-Difficult notes (M=4.41, SE=0.18), F(1, 117)=4.95, p=.028, η_p^2=.04, d=.46. There was no main effect of note format, F(1, 117)=1.87, p=.174, η_p^2=.02, d=.26, or interaction, $F < 1$ (Figure 3; Table 3).

Helpful. Students perceived using Less-Difficult notes (M=4.30, SE=0.19) as more helpful than More-Difficult notes (M=3.50, SE=0.18), F(1, 117)=8.92, p=.003, η_p^2=.07, d=.63. There was no main effect of note format or interaction, Fs < 1 (Figure 3; Table 3).

Discussion

The current experiment addressed a theoretically driven and practical question about learning and memory: How do note-taking format and level of difficulty impact learning from instructor-provided lecture notes? Our findings provide additional evidence in support of the material-appropriate difficulty framework (McDaniel & Butler, 2011; McDaniel & Einstein, 1989, 2005; McDaniel et al., 1988). Increasing note-taking difficulty (by including more blanks) led to better memory outcomes (free recall and inference)—but only in the outline notes condition. When giving a descriptive, fact-based science lecture, outline notes enhanced the ability to infer relations among the ideas presented. Verbatim short answer accuracy, which assesses memory for specific facts, was equivalent between outline and cloze note formats. Thus, by engaging in both item-specific and relational processing, students given outline notes were able to retain both types of information. In contrast, cloze notes appear to have induced only item-specific processing, supporting memory for specific facts alone.

This study also included several process measures, adding metacognitive and objective cognitive load measures to the material-appropriate difficulty literature. Despite overall better memory, outline notes were rated as more difficult and led to greater online cognitive load. In addition, notes in the more-difficult conditions were rated as more difficult, less enjoyable, less helpful, and as leading to lower comprehension than notes in the less-difficult conditions. More-Difficult Outline notes were also the least complete, but they improved free recall and inference accuracy compared to the other three note-taking formats. These results mirror other findings demonstrating that what students find easier, more helpful, and more enjoyable do not always match what most benefits their learning (Rohrer & Pashler, 2010).

These results suggest that outline notes and greater difficulty posed the greatest challenge for students. Yet, these conditions increased "desirable difficulty," or generative processes that support learning (Bjork & Bjork, 2011). Even though material-appropriate difficulty was intended to be achieved in both More-Difficult Outline and Less-Difficult Outline conditions, increasing the difficulty of outline notes required students to be more actively involved in making sense of the information, which led to better learning. Moreover, difficulty alone was not sufficient to increase learning—learning was not improved in the More-Difficult Cloze notes condition. Material-appropriate processing was required in order for difficulty to be desirable (McDaniel & Butler, 2011).

By showing better learning in the More-Difficult Outline notes condition, these findings do not align with cognitive load theory (Paas, 1992; Sweller et al., 1998). Providing instructor notes is often intended to reduce students' mental effort during lecture, to increase their ability to process the information more deeply and ask questions (Austin et al., 2002). Yet, we found that increasing difficulty was more beneficial, and modifying difficulty only mattered with one of the two note-taking activities. It remains likely that further increasing the difficulty of outline notes would increase cognitive load to such an extent that material-appropriate processing would no longer have a benefit (Einstein et al., 1990). It is also possible that increasing the difficulty of cloze notes would lead to even lower learning.

Limitations

The conditions created for the current experiments were not intended to fully replicate the conditions found in the classroom, and some limitations must be noted. For example, the lecture was brief, informationally dense, and delivered at a relatively quick pace, which may not be representative of many lectures. Furthermore, students were tested within an hour of listening to the lecture, so these results cannot speak to whether the effects extend over longer retention intervals. This study specifically examined the encoding benefit of note-taking, and therefore any impact of reviewing the notes is unknown.

In addition, we used a descriptive lecture, and thus our results cannot speak to other lecture types (i.e., a narrative lecture). Indeed, the nature of a lecture (i.e., descriptive versus narrative) is not always likely to be straightforward. The same lecture could be delivered in multiple ways—as a story that emphasises the relational aspects of the information, or without making the connections between ideas explicit and thus emphasising the item-specific aspects of the information. Furthermore, the same lecture could switch between emphasising the two types of semantic processing. Cook and Mayer (1988) note that, without extensive tailored training, students often have difficulty identifying organisational structures (i.e., relational information) of descriptive prose, even though the prose contains both item-specific and relational information.

Although we suggest that increasing difficulty in the More-Difficult Outline notes condition increased generative processing, an alternative possibility is that the Less-Difficult Outline notes no longer induced relational processing. Specifically, students may not have processed the relational information because they did not organise the lecture content on their own. Instead, they simply filled in missing words. Thus, Less-Difficult Outline notes may have induced processing that is more similar to the cloze notes (i.e., item-specific processing) than to the More-Difficult Outline notes. Again, this possibility demonstrates the importance of considering the level of difficulty in encoding tasks. Reducing difficulty may either decrease generative processing or alter the nature of semantic processing entirely. More research is needed to test these possibilities.

Conclusion

The current experiment provides strong evidence that both note format and level of difficulty are important variables that impact the efficacy of guided notes. These findings also replicate and extend the material-appropriate difficulty framework from self-paced (text-based) to instructor-paced (audio-based) learning materials. These findings support a recommendation for instructors and students: when learning from descriptive lectures, students should take notes using outlines that require a greater degree of generative processing. More generally, instructor-provided notes should require generative processing, even if students experience difficulty. This point is particularly relevant considering our samples' reports that instructors were much more likely to provide complete than partial notes. Furthermore, only a third of students stated that they elect to create outlines during science lectures, suggesting that many students do not choose the optimal note format. By adopting a "less is more" approach to instructor-provided outline notes, students will be encouraged to assume a more optimal method of processing descriptive lecture information. Given the ubiquitous use of lectures and the importance of note-taking for capitalising on this learning opportunity, the development of outline notes for descriptive lectures is a promising educational intervention.

Acknowledgements

The authors thank Cara H. Cashon, Keith B. Lyle, John R. Pani, and Kate E. Snyder for helpful feedback on this research, and Mark McDaniel for feedback on the manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This article is based on a PhD dissertation submitted to the University of Louisville. The research was supported by a Dissertation Research Award from the Graduate Network in Arts and Sciences at the University of Louisville.

ORCID iD

Marci S DeCaro https://orcid.org/0000-0001-6753-0725

Note

1. The study was not designed to examine interactions with these individual differences, thus the sample size was insufficient to test additional interactions.

References

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

Austin, J. L., Lee, M. G., Thibeault, M. D., Carr, J. E., & Bailey, J. S. (2002). Effects of guided notes on university students' responding and recall of information. *Journal of Behavioral Education*, *11*, 243–254.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York, NY: Worth Publishers.

Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, *106*, 849–858.

Boyle, J. R. (2012). Note-taking and secondary students with learning disabilities: Challenges and solutions. *Learning Disabilities Research & Practice*, *27*, 90–101.

Bui, D. C., & Myerson, J. (2014). The role of working memory abilities in lecture note-taking. *Learning and Individual Differences*, *33*, 12–22.

Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology*, *105*, 299–309.

Chen, X. (2002). *Teaching undergraduates in U.S. postsecondary institutions: Fall 1998*. Washington, DC: Government Printing Office (NCES Publication No. 2002-209). Retrieved from http://nces.ed.gov/pubs2002/2002209.pdf

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

Cook, L. K., & Mayer, R. E. (1988). Teaching readers about the structure of scientific text. *Journal of Educational Psychology*, *80*, 448–456.

Cornelius, T. L., & Owen-DeSchryver, J. (2008). Differential effects of full and partial notes on learning outcomes and attendance. *Teaching of Psychology*, *35*, 6–12.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*, 223–234.

DiVesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology*, *63*, 8–14.

Dunkel, P., & Davy, S. (1989). The heuristic of lecture note-taking: Perceptions of American & international students regarding the value & practice of notetaking. *English for Specific Purposes*, *8*, 33–50.

Einstein, G. O., & Hunt, R. R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 588–598.

Einstein, G. O., McDaniel, M. A., Bowers, C. A., & Stevens, D. T. (1984). Memory for prose: The influence of relational and proposition-specific processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 133–143.

Einstein, G. O., McDaniel, M. A., Owen, P. D., & Cote, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory and Language*, *29*, 566–581.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*, 717–741.

Hunt, R. R. (2003). Two contributions of distinctive processing to accurate memory. *Journal of Memory and Language*, *48*, 811–825.

Hunt, R. R. (2013). Precision in memory through distinctive processing. *Current Directions in Psychological Science*, *22*, 10–15.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, *20*, 497–514.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 429–446). Hillsdale, NJ: Lawrence Erlbaum.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*, 772–775.

Karpicke, J. D., & Roediger, H. L., III. (2010). Is expanding retrieval a superior method for learning text materials. *Memory & Cognition*, *38*, 116–124.

Katayama, A. D., & Robinson, D. H. (2000). Getting students "partially" involved in note-taking using graphic organizers. *The Journal of Experimental Education*, *68*, 119–133.

Kiewra, K. A. (1985a). Investigating notetaking and review: A depth of processing alternative. *Educational Psychologist*, *20*, 23–32.

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

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

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

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

Mayer, R. E. (1985). Structural analysis of science prose: Can we increase problem-solving performance?. In B. K. Britton & J. B. Black (Eds.), *Understanding expository prose* (pp. 65–87). Hillsdale, NJ: Lawrence Erlbaum.

Mayer, R. E. (1987). Instructional variables that influence cognitive processing during reading. In B. K. Britton & S. Glynn (Eds.), *Executive control processes in reading* (pp. 201–216). Hillsdale, NJ: Lawrence Erlbaum.

McDaniel, M. A., & Butler, A. C. (2011). A contextual framework for understanding when difficulties are desirable. In

A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 175–198). New York, NY: Psychology Press.

McDaniel, M. A., & Einstein, G. O. (1989). Material appropriate processing: A contextualistic approach to reading and studying strategies. *Educational Psychology Review*, *1*, 113–145.

McDaniel, M. A., & Einstein, G. O. (2005). Material appropriate difficulty: A framework for determining when difficulty is desirable for improving learning. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 73–85). Washington, DC: American Psychological Association.

McDaniel, M. A., Einstein, G. O., Dunay, P. K., & Cobb, R. E. (1986). Encoding difficulty and memory: Toward a unifying theory. *Journal of Memory and Language*, *25*, 545–656.

McDaniel, M. A., Einstein, G. O., & Lollis, T. (1988). Qualitative and quantitative considerations in encoding difficulty effects. *Memory & Cognition*, *16*, 8–14.

Mitchell, D. B., & Hunt, R. R. (1989). How much “effort” should be devoted to memory? *Memory & Cognition*, *17*, 337–348.

Morgan, C. H., Lilley, J. D., & Boreham, N. C. (1988). Learning from lectures: The effect of varying the detail in lecture handouts on note-taking and recall. *Applied Cognitive Psychology*, *2*, 115–122.

Neef, N. A., McCord, B. E., & Ferreri, S. J. (2006). Effects of guided notes versus completed notes during lectures on college students’ quiz performance. *Journal of Applied Behavior Analysis*, *39*, 123–130.

Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*, 429–434.

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

Ponce, H. R., & Mayer, R. E. (2014). An eye movement analysis of highlighting and graphic organizer study aids for learning from expository text. *Computers in Human Behavior*, *41*, 21–32.

Quintus, L., Borr, M., Duffield, S., Napoleon, L., & Welch, A. (2012). The impact of the Cornell note-taking method on students’ performance in a high school family and consumer sciences class. *Journal of Family & Consumer Sciences Education*, *30*, 27–38.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, *28*, 164–171.

Richland, L. E., Bjork, R. A., Finley, J. R., & Linn, M. C. (2005). Linking cognitive science to education: Generation and interleaving effects. *Proceedings of the Cognitive Science Society*, *27*, 1850–1855.

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher*, *39*, 406–412.

Svinicki, M., & McKeachie, W. J. (2011). How to make lectures more effective. In W. J. McKeachie (Ed.), *McKeachie’s teaching tips: Strategies, research, and theory for college and university teachers* (pp. 55–71). Belmont, CA: Wadsworth.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.

Van Merrienboer, J. J. G., & Ayres, P. (2005). Research on cognitive load theory and its design implications for e-learning. *Educational Technology Research and Development*, *53*, 5–13.

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

Williams, R. L., & Eggert, A. C. (2002). Notetaking in college classes: Student patterns and instructional strategies. *The Journal of General Education*, *51*, 173–199.

Williams, W. L., Weil, T. M., & Porter, J. C. K. (2012). The relative effects of traditional lectures and guided notes lectures on university student test scores. *The Behavior Analyst Today*, *13*, 12–16.