The Effect of a Break on Textbook Reading Comprehension

Robert Eastwood
Robert.Eastwood@Colorado.EDU

Follow this and additional works at: https://scholar.colorado.edu/honr_theses

Part of the Cognitive Psychology Commons

Recommended Citation
https://scholar.colorado.edu/honr_theses/1500

This Thesis is brought to you for free and open access by Honors Program at CU Scholar. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
The Effect of a Break on Textbook Reading Comprehension

R. Mason Eastwood

University of Colorado Boulder

Thesis Advisor: Alice F. Healy, Psychology and Neuroscience

Honors Council Representative: Eliana Colunga, Psychology and Neuroscience

Outside Member: Rolf Norgaard, Program for Writing and Rhetoric

Date of defense: October 24, 2017

Author Note

R. Mason Eastwood, Department of Psychology and Neuroscience, University of Colorado Boulder.

Special thanks goes to Alice Healy for her consistent advising on this project, and for her always being willing to check that data analysis was done properly. Special thanks also goes to Adam Young for his immense help in designing and manifesting the experiment in its computerized form. I would also like to thank Akira Miyake for his remarks on this work after hearing it presented in his lab.
Abstract

When, if ever, is it beneficial to take a break from intellectual work or study? Here begins an investigation into the broad, important, and scientifically untapped issue of the benefit of break taking by placing subjects into one of two experimental conditions: (a) those who took a 5-min break during the reading of a college level astronomy textbook chapter, or (b) those who did not take a break. The effect of the break was quantified by subjects’ score on two multiple-choice tests designed to measure comprehension of the material contained in the textbook chapter, and by the amount of time it took subjects in each condition to read the chapter. One multiple choice test was administered immediately after subjects read the chapter, and the other 48 hrs later. An equal number of questions on each test corresponded either to the section of the text read prior to the break, or to the section of the text read after the break. Reading time was also broken down by the same two sections of the text. There was no effect of the break on either reading comprehension or reading time. In other words, the break neither helped nor harmed subjects’ reading comprehension, and neither increased nor decreased subjects’ reading speed.

Nevertheless, strong conclusions about the effectiveness or lack thereof of breaks from intellectual work cannot be drawn from this novel investigation. Instead, a future body of research should seek to build upon the groundwork that has been laid here.
The Effect of a Break on Textbook Reading Comprehension

There are clear physiological limits to the amount of time a person can spend engaged in academic study or intellectual work in a given period. At some point, people must eat, drink, or sleep in order to survive. Thus, it would be impossible to spend 24 consecutive hours reading scholarly articles or studying for an exam without being seriously distracted by fatigue, hunger, or thirst. However, there is no experimental consensus regarding how long people can study, read, or write before experiencing a noticeable decrease in productivity and efficiency for the task at hand within more reasonable temporal limits. Some questions will further illuminate the problem. Can a student read several chapters of academic material successively over a period of several hours? And can that student comprehend the material just as well as if the chapters were read separately in shorter intervals? If it would take 4 hrs to write a term paper, does that mean that the whole paper should be written within one 4-hr period? Is there a difference in students’ ability to comprehend a lecture, or a professor’s ability to deliver a lecture, if the lecture is 1 hr versus 5 hrs long?

There is a fundamental issue that runs through each of the above questions. When, if ever, is it beneficial to take a break from intellectual work? How long should the break be? How long should the intervals of study or work be? It is not hard to imagine that there are several variables at play here, many of which will be illuminated below. It is also not hard to see the practical importance of investigating these questions, and ones like them, for students and professionals alike. Consider two hypothetical students. One attempts to study for hours on end with no breaks, and believes she is being as productive as possible, learning as efficiently as possible. The other studies for only 20 or 30 minutes at a time, and takes equally long breaks between study intervals. He believes that he is replenishing depleted cognitive resources
necesary to maximally engage with the task at hand. Of course, both students may lie at opposite extremes of a study/break and benefit/detriment ratio continuum, but the question remains: All other things being equal, which student performs better?

In order to begin to experimentally investigate this novel research topic, it is necessary to arrive at some specific parameters. Without any prior evidence to consult, determining what type and duration of work or study to use, what kind of break to use, how long of a break to use, and how to test the effect of a break proved to be a challenge. Nevertheless, an experiment was designed whereby subjects read a full 20-page chapter of a college level astronomy textbook and were tested for comprehension both immediately after reading, and again two days later. Half of the subjects took a 5-min break roughly midway through reading the chapter, and then finished the chapter whereas the other half read it all the way through with no break. Thus, the following question was examined:

Is it better for comprehension to take a 5-min break while reading a textbook chapter, or to read the entire chapter with no break? In other words, will those who take a 5-min break while reading a textbook chapter perform better on a test of the material from the chapter than those who do not take a break? This experimental question contains within it questions regarding the nature of sustained attention to a task, the validity and applicability of cognitive resource theory, mental and physical fatigue, and an array of other subjects. Before delving into some of the possible reasons why taking a break during reading might be beneficial based on existing research areas, it is most important to reiterate that this basic experiment has not been done before. There is no line of research directly addressing breaks from continuous study, such as reading a textbook chapter, to draw upon.
An initial search led to the literature on the spacing effect. Subjects in these studies generally memorize verbal lists or associated word pairs either all at once (massed study), or in a broken up fashion (spaced study) by which study of one item is interrupted with study of another item, or nothing at all (Cepeda, Pasher, Vul, Wixted, & Rohrer, 2006). The literature on spaced vs. massed study is dense, but it seems that within certain temporal confines, spaced study results in higher performance on later testing for verbal memory tasks than massed study (Cepeda et al., 2006). There are, however, fundamental differences between spaced vs. massed study and taking a break while reading a textbook chapter that render the two topics incommensurable. Spacing the study of an item results in studying the same item two or more times (Cepeda et al., 2006). When reading a textbook chapter, the information studied before a break is not the same information that is studied after the break, albeit the material of the whole chapter is likely related to itself in some way, so the break does not result in seeing the same information twice. Furthermore, learning an entire chapter’s worth of information cannot be equated to learning an associated word pair, fact, list, or other paradigm commonly used in spacing effect experiments.

Like the spacing effect literature, all other areas of research and theorization that will be mentioned in this literature review are at best tangential to the central question of whether or not a break during textbook reading is beneficial. One such area is that of mind wandering (MW). Despite some controversy over the various ways to measure MW, which can be loosely defined as thinking about something other than the task at hand, research on it has become increasingly prevalent (Smallwood & Schooler, 2015). It has been shown that those who experience more MW, specifically those who think task unrelated thoughts (TUTs), while reading comprehend less of the material being read (Smallwood, 2011). The MW/comprehension tradeoff makes sense because if one is not thinking about what is being read, one is not effectively reading, and
will likely not comprehend what is being read. Investigations into what causes TUTs while reading reveal that interest in the material being read is a key moderating factor in the frequency of MW, more so than the perceived difficulty of the material being read (Giambra & Gradsky, 1989; Smallwood, Nind, & O’Connor, 2009). One possible benefit of taking a break during textbook reading is that it might reduce the occurrence of MW, especially if the reading material is uninteresting to the reader, by allowing the reader to refocus attention on the task at hand. Although it has been shown in media other than reading, such as online lectures, that breaking up learning with interpolated quizzes reduces MW, the effect of a pure break (not filled with study of previously encountered information) on MW frequency has never been tested (Spuznar, Khan, & Schacter, 2013).

Interestingly, the effects of interspersing quizzes throughout the presentation of material on comprehension of the material, not just MW, have been investigated. There is laboratory evidence that quizzing subjects multiple times during the presentation of information leads to better test performance at a later time, even on presented material that was not quizzed, than quizzing subjects only after all of the material to be learned has been presented (Healy, Jones, Lalchandani, & Tack, 2017). The fact that un-quizzed information is better recalled at test for the subjects who were quizzed multiple times during the presentation of material suggests that breaks in the presentation might have a beneficial effect over and above that of quizzing. Indeed, Healy et al. (2017) attribute the benefit of interspersed quizzing to an increase in task engagement for subjects in the interspersed condition, as measured by self-report survey regarding boredom, motivation, MW, and fatigue. If a break from textbook reading does increase comprehension of the text, then it may be because of a similar increase in task engagement as found by Healy et al. (2017).
MW may just be a symptom of a larger mechanism that causes a detriment to reading comprehension over time. An attractive candidate for that mechanism is the strength model of self-control, which was derived from research on ego depletion. The strength model purports that engaging in an act of self-control (e.g., inhibiting a response, performing a task despite the desire to do something else) depletes one’s ability to engage self-control (i.e., ego) on a subsequent task (Baumeister, Vohs, & Tice, 2007). The fundamental idea behind ego depletion and the strength model is that there is some limited mental resource which is drawn from when exerting self-control that must be replenished in order to operate at full cognitive capacity (Muraven & Baumeister, 2000). The relevance of the ego depletion literature to the potential benefit of taking a break while reading a textbook chapter is clear. If reading for a prolonged period requires self-control (e.g., suppressing TUTs), and if the reader’s ability to exert self-control is depleted before the chapter is finished, then readers who do not get a break should have less opportunity to replenish whatever resource they have been draining. Readers who get to take a break might be able to replenish the limited self-control resource, or at least stop exerting self-control for a period of time, which could result in greater attention to the material being read, and ultimately greater comprehension of that material.

Despite the attractive applicability of the strength model to the topic of taking a break while reading, there has been recent controversy regarding the validity and robustness of the effects found in the recently popularized body of ego depletion literature (Lurquin et al., 2016). It has been shown that one’s belief about whether or not cognitive resources are limited might trump any depleting effect of tasks commonly used in ego depletion studies (Job, Dweck, & Walton, 2010). In other words, someone who believes that mental resources are unlimited may outperform someone who believes the opposite. If the effect of ego depletion depends upon what
an individual believes about mental resource availability, then the strength model cannot serve as a definitive explanation for any detriment in self-control ability that occurs over time. More striking, however, is a meta-analysis on effect sizes, which concluded that the true effect size of ego-depletion studies may not be different from zero, and that this previously unnoticed lack of robustness likely stems from publication bias (Carter & McCulough, 2014). These discouraging findings regarding the validity of the strength model prompt further investigation into why a reader might need to take a break while reading a textbook chapter for comprehension.

Following the rationale provided by Healy et al. (2017), the relevance of fatigue and boredom will be considered.

The effect of rest breaks has been examined in the field of ergonomics. The focus of research on breaks in ergonomics is not to improve comprehension or learning, but to improve worker productivity. Nevertheless, there are likely parallels between increased productivity and increased reading comprehension. Frequent, short (less than 10-min) rest breaks from computer work have been shown to increase the amount of work that can be done in an hour, and the well-being of workers, with the benefits attributed to decreased physical and mental fatigue (Henning, Jacques, Kissel, & Sullivan, 1997). Although the concept of mental fatigue was not well defined, the takeaway from this finding, and other studies of breaks from relatively menial work (compared to comprehending the subject matter of a college level textbook), is that physical and mental fatigue might play a role in diverting attention from the task at hand (Dababneh, Swanson, & Shell, 2001).

It has even been posited that the previously discussed, questionable evidence on ego depletion might be remedied by investigating physical fatigue; the resource pool that is drawn from when exerting self-control might be analogous to, or even the same as that of physical
fatigue, which has been studied both psychologically and biologically (Evans, Boggero, & Segerstrom, 2015). A simpler explanation, not relying on any central, limited cognitive resource pool, pertaining to textbook reading comprehension is that sitting in a chair and focusing on something for a long period of time might result in physical discomfort that could be alleviated with a break in which the reader can not only stop exerting mental effort and thus become more mentally comfortable, but also get up from the chair to stretch.

In addition to alleviating mental or physical fatigue, a break from textbook reading might reduce boredom. Although an operational definition of boredom does not exist in the psychological literature, a functional one is lack of interest or motivation to engage in the task at hand. There has been qualitative research on the phenomenon of boredom suggesting that boredom can arise as a result of either overly high, or overly low cognitive arousal (Martin, Sadlow, & Stew, 2006). If the reader finds the material of the textbook chapter uninteresting, and is not in some way stimulated by it, then boredom may occur. Conversely, if the reader finds the material too difficult to comprehend at first glance, then boredom might occur. In concordance with the above reasoning, Kole, Healy, and Bourne (2008) found that interrupting an easy, monotonous data entry task with a more challenging one (e.g., alternating the concluding keystroke rather than using a single concluding keystroke) resulted in increased performance on the data entry task. Specifically, the addition of a more challenging task in the Kole et al. (2008) experiment eliminated the well documented speed accuracy tradeoff, whereby speed of data entry increases and accuracy decreases over time (Healy, Kole, Buck-Gengler, & Bourne, 2004). The addition of the more challenging task was proposed as a cognitive antidote to the detrimental effects of boredom and fatigue that arise during the data entry task (Kole et al., 2008).
Another qualitative study found that the ways that college students in an educational setting deal with boredom often include disengaging with the content to be learned, which would likely cause a comprehension detriment (Daniels, Tze, & Goetz, 2015). It follows that boredom may occur for readers while reading an entire textbook chapter, resulting in less comprehension of the material presented in the chapter. It might also be that taking a break while reading the chapter could serve as a cognitive antidote to symptoms of boredom, allowing readers to comprehend more of what they read.

A final reason why taking a break might be beneficial to textbook chapter readers has less to do with symptoms of fatigue or discomfort, and more to do with memory. It has been shown that engaging in 5 min of wakeful resting, which can be defined as sitting quietly without engaging in any task or activity, after reading a passage results in better memory of the passage, as measured by a test 7 days later, than engaging in a relatively cognitively demanding task after reading (Dewar, Alber, Butler, Cowen, & Della Sala, 2012). The explanation for the positive effect of wakeful resting is that readers consolidate more of the material they read by sitting quietly, even if they are not actively trying to do so, into long term memory than if they immediately occupy their minds with other matters (Dewar et al, 2012). Although the break for the purposes of the current experiment allows break takers to engage in tasks of their choosing, some effect of wakeful resting might occur. If there is indeed an effect of wakeful resting for those readers who take a break, it follows that any benefit would only be seen for the section of the text read prior to the break because one cannot possibly consolidate information into long term memory that one has not yet encountered.

It is hard to state whether or not a break will be beneficial in this specific and rather arbitrarily arrived at set of experimental parameters, especially considering the wide and
disconnected breadth of literature that had to be consulted due to the absence of any more
directly relevant evidence from which to draw upon. From the research domains that have been
surveyed, it seems that several factors could mediate the effect, if any, of a break from study.
These include task difficulty, task duration, interest in the task, motivation to comprehend the
material, and several others. Many of these potential mediators were considered in the design of
the experiment that is described below. It is hypothesized that the 5-min break will benefit
reading comprehension due to some combination of decreased mental fatigue, increased
motivation, decreased boredom, and decreased mind wandering.

Method

Subjects

Sixty introductory psychology undergraduate students from the University of Colorado
Boulder completed this two-session experiment in exchange for partial course credit. A number
of additional students participated in Session 1 of the experiment, but were removed from data
analysis due to not returning for Session 2 of the experiment, or due to a malfunction of the
computer program used to administer the experiment.

Materials

Subjects read the first chapter of *The Cosmic Perspective*, written by Bennet, Donahue,
Schneider, and Voit (2014), which is a college-level astronomy textbook. Each page of the
chapter was transferred from an online version of the textbook via a photo editing software into a
computer program written with Python and PsycoPi. With the photo editing software,
supplementary mathematical text boxes and extraneous informational text boxes were removed
from the chapter so that only the 20 pages of main text were presented. Subjects saw one page at
a time, and a timer was built into the program so that the amount of time spent reading each page
could be measured. Additionally, subjects had to spend at least 10 s viewing each page before they could move to the next, and once they moved to the next, no previous pages could be accessed. The chapter contained four subsections, which were clearly marked with bold headings. The first two subsections will hereafter be referred to as “Section 1,” and the second two subsections as “Section 2.” Section 1 contained 13 pages, and Section 2 contained 7 pages. A considerable challenge in designing the experiment was deciding whether or not to make the amount of text read prior to and after the break equal. Instead, we decided to place the break between subsections such that a reasonable stopping place was provided because it seems more likely that students who take study breaks in their real lives would not choose to stop reading in the middle of a paragraph or sentence. Furthermore, if the necessity of a break depends upon fatigue and/or boredom, then having the section read prior to the break be longer than the section read after the break should help maximize such effects.

Two multiple-choice tests were constructed (Test X and Test Y) so that subjects could be tested for comprehension of the chapter both immediately following reading and 48 hours later. Both tests were administered within the same program that presented the textbook chapter pages. Each test contained 12 questions, and each question contained one correct answer and three foils so that the correct answer was “A,” “B,” “C,” or “D.” Half of the test questions and answer choices were written by the authors of the textbook, and I wrote the other half, such that each test contained six questions written by the textbook authors and six by myself. The reason for my writing half of the questions was that the textbook did not provide a sufficient number. Additionally, half of the questions on each test pertained to Section 1 of the chapter, and half of the questions to Section 2. Finally, half of the questions from each Section on each test were factual, and half were conceptual. Factual questions were those to which the answer came
directly from the body of text, and conceptual questions were those to which additional synthesis of knowledge gleaned from the text had to be applied. Thus, both tests contained an equal number of questions belonging to one of eight categories: self-written conceptual from Section 1, self-written factual from Section 1, textbook-derived conceptual from Section 1, textbook derived factual from Section 1, and again for all mentioned but from Section 2. See Appendix A to view both tests.

Following the immediate posttest in the first session, subjects responded to several questionnaire items regarding motivation to succeed on the posttest, interest in the chapter material, degree to which their minds drifted from the subject matter, perceived difficulty of the posttest, and boredom. Each questionnaire item about subjects’ experience reading the chapter was assessed twice; once for Section 1 and once for Section 2. Following the posttest after the 48-hr retention interval, subjects again answered questionnaire items about the posttest, and also answered a question about whether or not they had spent any time engaged with astronomy related material during the retention interval. All questions were rated on a scale from one to seven, one being the least prior knowledge, motivation, etc., and seven being the most. See Appendix B to view all questionnaire items.

For subjects in the break condition a 5-min period without the possibility to read the chapter was inserted between Sections 1 and 2. A simple computer game was programmed into that period for subjects to play if they chose to, and there were also non-astronomy related books and journals for subjects to read that were made available. Due to the fact that the experimenter was not in the same room as subjects during the experiment, the activity chosen by subjects in the break conditions was not explicitly monitored (see the Discussion section for consideration of this limitation).
Design

A 2x2x2 mixed factorial design was employed with the between-subjects variable of break condition (break vs. no break), and the within-subject variables of section of the chapter (Section 1 vs. Section 2) and retention interval (immediate vs. 48 hrs after reading). The dependent variables were time spent reading (each page had a timer embedded in it) and proportion correct on the posttest. The former dependent variable was included in order to determine if there was any effect of the break on time spent reading. It might be that those who took a break either read the second section faster or slower than those who did not. Subjects were randomly assigned to one of the break conditions, and test order was fully counterbalanced so that half of the subjects in each break condition took test X first, and half took test Y first.

Procedure

In an effort to make the experience maximally similar for every participant, and to have a method of data collection less prone to error, we decided to computerize everything. As a result, there was very little interaction by subjects with the experimenter, except for the signing of the consent form, and the initial greeting.

Subjects came into the lab, and were instructed to go into one of four rooms with a computer in it after placing their phone in a basket located in the main area of the lab. The experimenter came into the room with the participant, initiated the program, and told the subjects to follow all of the on screen instructions. The experimenter then closed the door for the remainder of Session 1 of the experiment.

All subjects received the same set of instructions on the computer with the exception of a couple of sentences specific to the break condition. All subjects were told that they would read the textbook chapter, be tested on it, and fill out a short questionnaire. Those in the break
condition were additionally told that they would be taking a 5-min break after completing the second of four subsections of the textbook chapter. We decided to let subjects in the break condition know they would be taking a break rather than surprise them with it in the middle of the chapter because doing so made more sense when thinking about the study habits of students; it seems more likely that students who take breaks on their own accord would be aware of their plan to take a break before the break occurred than that students who take breaks do so without thinking about it or knowing it will happen.

Each page of text required subjects to read for at least 10 s before pressing a button to continue to the next page. These 10-s periods were a safeguard against anyone who might have otherwise rushed through the experiment and provided unusable data. One page of text was presented at a time, and total time spent on each page was recorded with a timer built into the computer program.

When subjects in the break condition reached the end of Section 1 of the chapter, they saw a screen that told them they would take a 5-min break. The break was timed so that every participant in the break condition took a break lasting exactly 5 min. On the break screen was a computer game (space invaders) built into the experiment program, which subjects could play. Subjects were also told that they could find magazines, romance novels, and psychology journals to read nearby if they wanted to read them. A challenge in designing activities for subjects to engage in during the break stemmed from our desire to give them the freedom to do what they might normally do during a study break without allowing them access to the internet, on which they might look up astronomy related material. We also did not want them to have their phones, even though many students likely use their phones during study breaks. Smartphones not only pose a potential distraction for subjects while they were not supposed to be taking a break, but
they also provide access to the internet, and those two detriments seemed to outweigh the benefit of ecological validity.

When subjects in both conditions completed the chapter, there was a 5-min interval before the comprehension test was administered in which demographic information was collected, and the bulk of the questionnaire was filled out. This 5-min interval, which is distinct from the break and was applied to subjects in both break conditions, provided a time buffer aimed at lowering the amount of material from Section 2 of the chapter that might linger in working memory and cause a disproportionate number of questions regarding Section 2 of the chapter to be answered correctly (Baddeley & Hitch 1974). To clarify, students are not usually tested on chapter material as soon as the last word is read, and the 5-min post-reading interval was included in order to help mitigate the possibility that working memory would add noise to the results.

Following the 5-min interval that occurred after reading, subjects proceeded to take the multiple choice test (either X or Y) corresponding to the counterbalancing scheme, after which they answered the final questionnaire items pertaining to their experience on the test itself.

In order to determine whether or not the effect, if one exists, of taking a break on textbook reading comprehension holds up over time, we had subjects return to the lab 48-hrs after their first session to take the second multiple choice test. At Session 2, subjects came into the lab, went to the same room they were in during Session 1 after giving up their phones, took the test, and then responded to questionnaire items dealing with their experience of the second test and with the amount of astronomy related material they encountered during the retention interval. Subjects were then debriefed.
Results

A mixed factorial analysis of variance (ANOVA) on the proportion of correct responses on the multiple choice tests, with the between-subjects factor of condition (break vs. no break) and the within-subjects factors of retention interval (Session 1 vs. Session 2) and section of text (Section 1 vs. Section 2), did not reveal a main effect of condition; there was no difference, $F(1, 58) < 1$, in posttest score between those in the break and those in the no break condition (see Figure 1). The only significant result from this analysis was a main effect of section, $F(1, 58) = 11.05, MSE = 0.363, p = .002$, such that subjects scored lower on test questions pertaining to Section 2 (.637) relative to Section 1 (.715) of the textbook chapter. Lower scores on Section 2 may either indicate that the subject matter and/or test questions were more difficult for Section 2, or that some combination of fatigue, boredom, MW, or lack of motivation set in for all subjects regardless of break condition.

In order to determine whether the break had any effect on time spent reading the textbook chapter, reading times for both Section 1 and Section 2 were also included as dependent variables in the above mentioned ANOVA (without the variable of retention interval). There was no main effect of break condition, $F(1, 58) < 1$ (break: 35.98 min, no break: 36.51 min), and there was no interaction between break condition and section, $F(1, 58) < 1$. In other words, there were no differences between break conditions with regard to reading time, even when analyzed by section.

A second mixed factorial ANOVA identical to the previous one with the addition of counterbalance condition (Test X first vs. Test Y first) as a between subjects factor was conducted in order to determine whether the two multiple choice tests were of equal difficulty. The significant interaction of session with counterbalance condition, $F(1, 56) = 14.25, MSE = $
0.54, \( p < .001 \), revealed that Test X (0.636 at Session 1, 0.622 at Session 2) did indeed yield lower scores than Test Y (0.739 at Session 1, 0.708 at Session 2). Notice that the mean scores for Test X were lower than those for Test Y regardless of whether the test was taken immediately after reading, or 48 hrs later. Furthermore, the significant three-way interaction of session, counterbalance condition, and section, \( F(1, 56) = 20.40, MSE = .445, p < .001 \), revealed that the difference in scores on each of the multiple choice tests was concentrated in questions pertaining to Section 2 of the textbook chapter such that those questions were more difficult on Test X than on Test Y (see Figure 2). Importantly, however, the addition of counterbalance condition to the analysis did not alter the significance of any effect of break condition.

The addition of question type (factual vs. conceptual) as a within-subject factor to the above ANOVA and the restriction to the immediate test (i.e., in Session 1) added insight into the issue of which multiple-choice questions on Test X were more difficult than on Test Y. The restriction of the analysis to the immediate test allowed the factor of counterbalance condition to effectively render a direct comparison between Test X and Test Y. There was a significant interaction of question type and counterbalance condition, \( F(1, 56) = 14.71, MSE = .389, p < .001 \), which suggests that conceptual questions were more difficult on Test X than on Test Y (see Figure 3). Interestingly, there was also a main effect of question type, \( F(1, 56) = 52.94, MSE = 1.40, p < .001 \), such that factual questions (0.611) were, in general, more difficult than conceptual questions (0.764). A simple summary of this set of results is that while factual questions were more difficult across both tests, conceptual questions on Test X were more difficult than conceptual questions on Test Y. Despite the significant three-way interaction of counterbalance condition, session, and section from the analysis covered in the above paragraph, there was no significant three-way interaction of counterbalance condition, section, and question type, \( F(1, \)
56) = 1.843, \( MSE = .078, \ p = 0.18 \), in the present analysis. Nevertheless, it is plausible that the difference in scores on the two multiple choice tests was due to more difficult conceptual questions from Section 2 on Test X than on Test Y.

One-way ANOVAs conducted on two questionnaire items, namely those measuring interest in the subject matter, and how much information from the chapter subjects knew prior to reading it, suggest that there may have been a failure in randomization. There was a marginally significant difference, \( F(1, 58) = 3.01, \ MSE = 8.067, \ p = .088 \), between conditions with regard to interest in the subject matter such that those in the no break condition were more interested, as measured on a 7-point scale, in the material (3.70) than those in the break condition (2.97). Additionally, there was a marginally significant difference, \( F(1, 58) = 3.35, \ MSE = 6.02, \ p = .072 \), between conditions with regard to how much information they knew from the chapter prior to reading it such that those in the no break condition claimed to know more (2.93) than those in the break condition (2.30). There were no differences between conditions with regard to any of the other questionnaire items, namely boredom, MW, motivation to comprehend the material, and perceived difficulty of the posttests.

A correlation matrix was produced with all of the dependent measures, separately for Sessions 1 and 2, as well as the questionnaire items as input. The reason for the production of the matrix was to get a sense of the predictive validity of the questionnaire items, which was a necessary endeavor considering the fact that all of the items were written exclusively for the present experiment, and not already established in the literature. The logic for determining the predictive validity of a questionnaire item for the purposes of the present experiment is as follows: If two items that should intuitively be correlated in a certain direction are indeed significantly correlated in that direction, then the item measured what it was intended to measure.
to some degree of accuracy. A few notable findings will be mentioned here. MW during Section 1 was positively correlated with boredom during Section 1, \( r(58) = .588, p < .001 \). This result may suggest that there is some phenomenological similarity between MW and boredom. Interest in the chapter material was negatively correlated with boredom for Section 1, \( r(58) = -.571, p < .001 \). Both of these correlations also held for Section 2 of the chapter: MW during the Section 2 was positively correlated with boredom during Section 2, \( r(58) = .508, p < .001 \), and interest in the chapter material was negatively correlated with boredom for Section 2, \( r(58) = -.679, p < .001 \). Interest in the chapter material was also negatively correlated with perceived difficulty of the posttest, \( r(58) = -.461, p < .001 \), and positively correlated with proportion correct on the immediate posttest, \( r(58) = .388, p = .003 \). Because none of these correlations were of the opposite direction as would be expected, and because they were significant, it seems reasonable to conclude that the questionnaire items constructed for the present experiment had some degree of predictive validity.

Discussion

Although these results do not support the hypothesis that taking a 5-min break during reading of a college level textbook chapter will benefit comprehension of the chapter, the fundamental questions raised above regarding the possible benefit of study breaks remain unanswered. It would be irresponsible to posit that the lack of differences in scores observed between subjects who either did or did not take a break during reading entailed that taking a break is positive, negative, or neutral with regard to reading comprehension. In other words, although the null results should be taken seriously in the context of these specific experimental parameters, this experiment cannot serve as grounds for either confirming or rejecting the broader hypotheses that a break from textbook reading might benefit reading time and
comprehension of the material, or that a break from any kind of intellectual work might increase productivity. This section will, then, be comprised of a discussion about this experiment’s limitations, and ultimately serve as a guide for designing future experiments aimed at investigating the novel experimental topic of the potential efficacy of breaks from periods of intellectual work.

The first unexpected limitation of the experiment was that subjects read the chapter much faster than was initially anticipated. Based upon evidence gathered from pilot versions of the experiment, it was assumed that subjects would spend approximately 1 hr reading the textbook chapter. However, the mean reading time for subjects in the current experiment was only 36 min. Although there is no literature to draw from which stipulates the temporal boundaries mediating any effect of a break on academic reading, 30 min does not seem long enough to ensure any benefit of a break on comprehension. In fact, it might even be possible that interrupting such a short period of reading with a break could actually serve more as a distractor than as a period for alleviated mental or physical fatigue, and thus would cause a detriment to reading comprehension. So, any follow-up experiment employing reading as the study medium should consider utilizing a longer sample of text for subjects to read.

The second unexpected limitation of the experiment was that there may have been a randomization failure. If randomization had been successful, there would not have been differences between conditions with regard to interest in the subject matter of the textbook chapter and information known about the material prior to reading the textbook chapter. Subjects in the no break condition were marginally more interested in the material than those in the break condition. Thus, if interest plays a role beneficial to comprehension over and above that of decreasing MW (Smallwood et al., 2009), such as mitigating boredom, then the effect of the
break may have been masked by the fact that those in the break condition were less interested in the material than those in the no break condition. It is also possible that those in the no break condition actually did experience more MW or boredom and thus have a comprehension decrement (Smallwood, 2011) relative to those in the break condition despite our not finding a difference due to the fact that we employed a simple questionnaire item to probe MW instead of the more rigorous measurements employed by the MW literature (Smallwood & Schooler, 2015). The other mechanism by which any beneficial effect of the break may have been masked is that those in the no break condition claimed to know marginally more information about the material prior to reading it. One interpretation of this difference in prior knowledge is that there would have been a difference in posttest score in favor of the break condition had prior knowledge been equal across groups. With these two potential randomization failures in mind, it might be beneficial for future experiments to employ a more detailed screening process assessing interest in the material to be read and general knowledge of the subject matter than was employed here.

Another limitation has to do with the number of multiple-choice questions administered on the posttests. There are at least two possible ways that the break might have benefited reading comprehension. The first is that it reduced some combination of MW, fatigue, or boredom, and thus provided a comprehension benefit for Section 2 of the chapter (Daniels et al., 2015; Evans et al., 2015; Smallwood, 2011). If alleviation of such symptoms were the function of the break, then Section 2 would benefit because both of the experimental conditions would experience the same buildup of MW, fatigue, or boredom during Section 1, but only the break condition would have those symptoms alleviated. The second is that the break provided an opportunity to more successfully rehearse material from Section 1 of the chapter (Dewar et al., 2012). Although these possibilities are not mutually exclusive, if one or the other was more powerful, then the effect of
a break on comprehension would be most evident in scores for test questions regarding either Section 1 or Section 2, but not both. With this logic in mind, it is clear that the current study did not employ enough multiple-choice questions in order to confidently detect any effect of a break on posttest score. Each multiple-choice test employed in the current experiment contained only six questions from Section 1 and six questions from Section 2. As a consequence, unless the sample size were increased, the differences in comprehension would have to be incredibly large for an effect to be detected in only six multiple-choice questions. Furthermore, even if the effect was somehow present across all 12 multiple-choice questions, there still might not be enough to detect it. So, a future experiment should employ far more multiple-choice questions that somehow cover different facts and concepts from the text in order to detect the effect, if any, of a break.

A potential valid criticism of the experiment is that we did not explicitly monitor the activities chosen by subjects in the break condition during the break. During the break, subjects had a choice between sitting quietly, playing space invaders on the computer, or reading romance novels, psychology journals, or magazines. The goal was to foster ecological validity while still maintaining control over subjects’ ability to get on the internet and potentially look up astronomy related information. However, it seems retrospectively that our logic was flawed. In an attempt to simulate subjects’ choice to take the kind of break they would in their natural environment, we effectively only gave them three options. And, although there is no evidence to support this claim, it became apparent to the experimenter that most subjects chose to play space invaders based upon the audible rapid key pressing (a requirement of successful gameplay) exhibited by most subjects in the break condition.
It may be that different kinds of breaks (e.g., wakeful resting vs. playing a computer game vs. taking a short walk) might have different effects on reading comprehension. Thus, in addition to explicitly monitoring subjects’ activity during a break, future research should consider assigning subjects to engage in specific, systematically varied activities during the break. For example, it might be detrimental to engage in a task that requires thought about something other than the text, such as playing a computer game, but beneficial to sit quietly and reflect on what has been read prior to the break. Furthermore, although we did ensure that all subjects in the break condition took an exactly 5-min break, it might be that breaks of different duration have different effects on comprehension. So, future research should consider systematically varying not only the content of the break, but also its duration.

In conclusion, although no effect of a break was found by this novel experiment, the potential for a new body of applied research has been born. Increasing productivity and efficiency is a noble goal for students and academic professionals alike, and determining the confines within which taking breaks is beneficial to that goal is undoubtedly important. It may be that there are vast individual differences in the ways that breaks do or do not benefit certain people, and it may be that the effectiveness of breaks is considerably different in different intellectual work environments. Nevertheless, for those interested in maximizing the efficiency and comprehension of textbook reading, this experiment certainly serves as a necessary starting point for illuminating the effect of a break on such measures. Due to the lack of prior research in this area, coming up with a set of experimental parameters here was akin to blindly shooting at a target the size of a paperclip in a football field. However, the stage has been set for future investigations to get closer to hitting that target.
References


characteristics and individual differences: A pre-registered study. *PLoS ONE, 11.*

doi:10.1371/journal.pone.0147770


Figure 1. Proportion correct on the multiple choice test immediately after reading (Day 1), and 48 hrs after reading (Day 2), broken down by break condition (break vs. no break), and section of the text to which the multiple choice questions pertained (Section 1 vs. Section 2). Error bars indicate standard errors of the mean.
Figure 2. Proportion correct on the multiple choice test immediately after reading (Day 1), and 48 hrs after reading (Day 2), broken down by section of the text to which the multiple choice questions pertained (Section 1 vs. Section 2), and counterbalance condition (XY vs. YX). XY refers to those who took test X during Day 1, and YX refers to those who took test Y during Day 1. Error bars indicate standard errors of the mean.
Figure 3. Proportion correct on the multiple choice for each counterbalance condition (XY vs. YX) immediately after reading, and for question type (conceptual vs. factual). Note that because only Session 1 is analyzed here, the breakdown by counterbalance condition effectively renders a comparison between each of the multiple choice tests, Test X and Test Y. Error bars indicate standard errors of the mean.
Appendix A

Test X

1. What makes up the majority of the mass in the Milky Way galaxy? (Factual, Section 1, Self-Written)
   - A) Stars
   - B) Planets
   - C) Dark Matter *
   - D) Intergalactic dust

2. Why do the patterns of the stars in our sky look the same from year to year? (Conceptual, Section 2)
   - A) Because the stars in the constellations all move at the same speeds and in the same directions, so they don’t change their relative positions.
   - B) Because the stars in the constellations are not moving.
   - C) Because the stars in the constellations move so slowly-typically about the speed of a snail—that their motions are not noticeable.
   - D) Because the stars in the constellations are so far away. *

3. How long does it take for our solar system to orbit the center of the Milky Way Galaxy? (Factual, Section 2, Self-Written)
   - A) Our solar system does not orbit the center of the galaxy
   - B) About 500,000 years
   - C) About 230 million years *
   - D) We cannot possibly calculate such a thing

4. Our solar system consists of ______. (Factual, Section 1)
   - A) the Sun and all the objects that orbit it *
   - B) the Sun and several nearby stars, as well as the planets and other objects that orbit these stars
   - C) the Sun and the planets, and nothing else
   - D) a few hundred billion stars, bound together by gravity

5. Which statement about motion in the universe is false? (Conceptual, Section 2)
   - A) Dark matter is the fastest-moving material in the universe. *
B) Your speed of rotation around Earth’s axis is faster if you live near the equator than if you live near the North Pole.

C) Some stars in the Milky Way Galaxy are moving toward us and others are moving away from us.

D) Except for a few nearby galaxies, all other galaxies are moving away from us.

6. Could we see a galaxy that is 20 billion light-years away? (Assume that we mean a “lookback time” of 20 billion years.) (Conceptual, Section 1)

A) Yes, we have already detected galaxies at that distance.

B) No, because a galaxy could not possibly be that far away.

C) No, because it would be beyond the bounds of our observable universe.

D) Yes, if we had a big enough telescope.

7. Approximately how many stars are in the Milky Way Galaxy? (Factual, Section 1, Self-Written)

A) Less than 1 million

B) More than 100 billion

C) Just over 1 billion

D) A few trillion

8. When we look at the Sun, are we looking back in time? (Conceptual, Section 1, Self-Written)

A) No, there is no such thing as looking back in time.

B) No, the Sun is close enough to us that its light reaches us immediately

C) Yes, the Earth’s atmosphere causes the light from the Sun to slow down as it reaches us, so depending our altitude we see the Sun as it was about 8 minutes ago.

D) Yes, sunlight takes about 8 minutes to reach us due to its distance from us, so when we observe the Sun, we see it as it was 8 minutes ago.

9. Why couldn’t we observe our solar system’s orbit around the center of the Milky Way galaxy with the naked eye even if we were positioned far above the galaxy? (Conceptual, Section 2, Self-Written)

A) The orbital speed of our solar system undetectably slow.

B) Even though we orbit the center of the Milky Way galaxy at around 500,000 mph, we are so far away from the center that it takes 230 million years to complete an orbit.

C) Our solar system does not actually orbit the center of the Milky Way galaxy.
D) Our solar system orbits the center of the Milky Way galaxy so fast that we could
never determine where it is at any given time without mathematical calculations.

10. A typical galaxy is a _____. (Factual, Section 1)

   A) relatively small, icy object orbiting a star
   B) collection of a few hundred million to a trillion or more stars, bound together by
      gravity *
   C) nearby object orbiting a planet
   D) system consisting of one or a few stars orbited by planets, moons, and smaller objects

11. What do astronomers mean by the “observable universe”? (Conceptual, Section 1, Self-
Written)

   A) The universe has only been expanding for a finite amount of time, so there is a point
      beyond which nothing exists.
   B) The observable universe is simply the part of the universe that has been seen already.
   C) All of the universe is technically “observable,” we just haven’t developed the
      technology to see it all yet.
   D) We can only see light that is less than 14 billion light years away because, due to the
      age of the universe, light from objects further than 14 billion light years away hasn’t had time to
      reach us yet. *

12. Consider a raisin cake expanding uniformly in an oven. Viewed from one of the raisins, you
would see ____. (Factual, Section 2)

   A) all other raisins moving away from you, with more distant raisins moving faster *
   B) all other raisins moving away from you at the same speed
   C) all other raisins moving away from you, with more distant raisins moving slower
   D) all raisins, including your own, growing in size as the cake expands

Test Y

1. Which is the best description of the motion of stars within the Milky Way Galaxy? (Factual,
Section 2, Self-Written)

   A) Stars move at random relative to each other, but all stars in the galaxy ultimately orbit
      around the galaxy’s center in the same direction. *
   B) All stars rotate around the center of the galaxy at the same speed
   C) All stars move in random directions at random speeds within the Milky Way Galaxy.
D) Stars do not move at all within the Milky Way Galaxy.

2. If astrology (horoscopes- not astronomy!) is based at least in part upon constellations, what is a logical reason for why it may be problematic? (Conceptual, Section 2, Self-Written)

   A) Due to the movement of the stars, the birth and death of stars, and the passing of time, the constellations that astrology utilizes will not exist at some point in the future.

   B) Constellations are permanent (stars don’t move relative to other stars), so that aspect of astrology cannot be scrutinized.

   C) Stars orbit the center of the galaxy at the same speed, regardless of their distance from the center, so even though they are all moving, the constellations will remain intact.

   D) Constellations, even when viewed from the exact same position on Earth, actually look very different on a nightly basis.

3. When was our solar system born? (Factual, Section 1, Self-Written)

   A) 14 billion years ago
   B) 6 million years ago
   C) 4.5 billion years ago
   D) 8.5 billion years ago

4. Astronomers infer that the universe is expanding because distant galaxies all appear to _____.
   (Conceptual, Section 2)

   A) rotate rapidly
   B) be made mostly of dark matter
   C) be growing in size moving away from us, with more distant ones moving faster
   D) be moving away from us, with more distant ones moving faster.

5. Which elements were present shortly after the birth of the universe, before any stars formed? (Factual, Section 1, Self-Written)

   A) All of the elements that exist today existed before the formation of stars
   B) Hydrogen only
   C) Helium only
   D) Hydrogen and helium

6. What is the ecliptic plane? (Factual, Section 2)
A) The plane of the Milky Way galaxy
B) The plane of Earth’s equator
C) The plane of Earth’s orbit around the Sun *
D) The plane of the Sun’s equator

7. When we look at an object that is 1,000 light-years away we see it _____. (Conceptual, Section 1)
   A) as it was 1,000 years ago*
   B) looking just the same as our ancestors would have seen it 1,000 years ago
   C) as it was 1,000 light years ago
   D) as it is right now, but it appears 1,000 times dimmer

8. When did humans first learn that the Earth is not the center of the universe? (Factual, Section 2)
   A) About 2,500 years ago
   B) Within the past 500 years *
   C) We haven’t; there is still considerable scientific debate about whether Earth is the center of the universe.
   D) About 1,000 years

9. Based on observations of the universal expansion, the age of the universe is about _____. (Factual, Section 1)
   A) 14 million years
   B) 14 trillion years
   C) 14,000 years
   D) 14 billion years *

10. Astronomically speaking, what does it mean to look back in time? (Conceptual, Section 1, Self-Written)
    A) The concept of looking back in time is a myth
    B) Light from distant objects takes a certain amount of time to reach us, so when we see that light, we see the object as it was at some moment in the past. *
    C) Astronomers don’t really look back in time, that’s just a phrase they use to describe the vastness of the universe.
D) There is no way to physically see the past, but astronomers infer the past state of distant objects with mathematical formulas, knowledge of the properties of light, and the laws of physics.

11. What is the best explanation for why galaxies grouped relatively close together, such as the Milky Way and the Andromeda Galaxies, DO NOT move away from one another, but instead are getting closer to each other? (Conceptual, Section 2, Self-Written)

   A) ALL galaxies are getting further apart from each other
   B) Astronomers are still puzzled by this phenomenon.
   C) They are close enough to have gravitational influence on each other. *
   D) These are mostly unknown interactions involving dark matter

12. In what sense are we “star stuff”? (Conceptual, Section 1)

   A) Movie stars and other people are all made of the same stuff, so we all have the potential to be famous.
   B) Nearly every atom from which we are made was once inside a star. *
   C) We could not survive without light from our star, the Sun.
   D) The overall chemical composition of our bodies is about the same as that of stars.
Appendix B

Questionnaire Items

The following questions were answered immediately after reading the chapter, with the exception of the last two. At session 2 of the experiment, subjects again answered questions regarding the posttest, and answered the last two questions.

How interested were you in the subject matter of the chapter? 1 is not at all interested, and 7 is very interested.

How motivated were you to comprehend the subject matter for the posttest? 1 is not at all motivated, and 7 is very motivated.

How difficult was the posttest? 1 is not at all difficult, and 7 is very difficult.

How often did you feel that your mind drifted away from the content of the chapter while reading the first two sections? 1 is never, and 7 is very often.

How often did you feel that your mind wandered away from the content of the chapter while reading the last two sections? 1 is never, and 7 is very often.

How much of this information did you already know? 1 is none, and 7 is all of it.

How bored were you while reading the first two sections of the chapter? 1 is not bored at all and 7 is extremely bored.

How bored were you while reading the last two sections of the chapter? 1 is not bored at all and 7 is extremely bored.

How much astronomy related information have you looked up, on the internet or in textbooks, since session one of this experiment? 1 is none, and 7 is a large amount.

How much have you thought about astronomy related material since session one of this experiment? 1 is not at all, and 7 is a lot of thought.