Metacognitive Prompts and the Paper vs. Screen Debate:
How Both Factors Influence Reading Behavior

A Thesis
Presented To
The Academic Faculty

By

Dar-Wei Chen

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science in Psychology

Georgia Institute of Technology

August, 2015

Metacognitive Prompts and the Paper vs. Screen Debate:
How Both Factors Influence Reading Behavior

Approved by:

Dr. Richard Catrambone, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Wendy Rogers
School of Psychology
Georgia Institute of Technology

Dr. Bruce Walker
School of Psychology
Georgia Institute of Technology

Date Approved: April 6, 2015

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

As online learning rises in popularity, students are increasingly learning through technology and without regular guidance from teachers. These learning environments differ from traditional classrooms in many ways and deliver different experiences. In this study, participants' learning environments were manipulated using two independent variables, each with two levels for a total of four conditions: study medium (text was presented either on paper or a screen) and prompt type (text was interspersed with prompts designed either to induce metacognitive processes or to be interacted with non-metacognitively). Ninety-two participants were each assigned to one of the four conditions in a between-subject design, read three expository texts, completed a comprehension test after each text, and responded to a survey at the end of the study. Participants who read text on paper tended to take more notes and spend more time studying than those who read from a screen, but performance was equal between the mediums. Participants receiving metacognitive prompts performed better than non-metacognitive participants on multiple-choice questions with an effect size comparable to those generated by educational interventions in existing literature; however, the performance difference was not statistically significant unless prompt response scores were controlled for. In addition, behavioral differences emerged between metacognitive participants (re-read more) and non-metacognitive participants (summarized more while reading). The results from this study can be used to inform dialogue about technology in classrooms and instructional design.

Keywords: Metacognition, self-regulation, instruction, learning technology, paper vs. screen, prompts, reading, learner engagement, motivation

CHAPTER 1

INTRODUCTION

Technology is now ubiquitous in many fundamental tasks of everyday life, including presentation of text for reading. According to The Council for Research Excellence (2009), American adults spend an average of 8.5 hours a day looking at a screen. On any given day, Americans are more likely to have read news online (39%) than in a newspaper (23%), despite a relatively robust 47% of Americans reading newspapers as recently as 2000 (Pew Research, 2012). Digital books are now outselling paper books on Amazon.com and digital textbooks saw a sales growth of 44.3% in 2011 (Schuetze, 2011). These and other statistics all point in the same direction: the electronic screen is quickly surpassing other forms of text presentation. At first glance, this technology transformation might seem unremarkable in the sense that people ostensibly will not be changing much in terms of executing the act of reading itself. However, as more people nowadays read on computers and mobile devices, researchers have become more interested in how (or if) reading changes when the text is presented on screens (as opposed to traditional paper sources).

For students in the classroom, changes in how they read can also stem from instructional design. For example, the use of embedded prompts in educational texts is a strategy that some educators use to facilitate learning. Improving education is hardly a cause that requires championing; therefore, researching study medium and prompts is important because of the potential effects that those two factors could have on the most basic and important education-related task of reading text.

1.1 Reading on paper and screens

Since computers started making their way into workplaces and homes in the late twentieth century, researchers have been studying the various cognitive effects of reading text on screens as opposed to traditional paper sources, and many of the findings have implications for learning. Current research suggests that paper presentation of text still has some advantages, even as people have become increasingly familiar with the backlight of a glowing screen. For example, people are slower when reading from a screen than from paper (Muter et al., 1982; Mayes, Sims, & Koonce, 2001). Skimming also takes more time on a screen than on paper (Muter & Maurutto, 1991). Even proofreading tasks are found to take longer for people using screens, and with less accurate performance, than those using paper (Creed, Dennis, & Newstead, 1987; Oliver, 1994; Wilkinson & Robinshaw, 1987).

Other measures indicate further advantages for paper texts. In a study by Wastlund et al. (2005), comprehension was found to be higher in paper readers than screen readers. Noyes, Garland, and Robbins (2004) found in their study that although overall reader comprehension was not significantly different between screen and paper conditions, screen readers endured higher cognitive workload to achieve that level of comprehension. This conclusion was corroborated by a related finding that screen readers reported feeling more tired and stressed than paper readers performing the same activities (Wastlund et al., 2005). On a subjective level, online learners and students still prefer print materials "for reasons of portability, dependability, flexibility, and ergonomics" (Spencer, 2006, p. 33) and people also prefer them when they need to study thoroughly (Jamali, Nicholas, & Rowlands, 2009; Buzzetto-More, Sweat-Guy, & Elobaid, 2007). These findings seem to indicate that for all of the advances made in screen technology, paper reading still "feels" better to most people.

However, the performance gap between screens and paper has generally been narrowing in recent years as new technologies address weaknesses in electronic reading (or as people generally become more familiar with technology). Along those lines, few broad conclusions can be drawn regarding the advantages of one medium over the other because many studies in the body of research contain findings that clash with each other, likely because of variations in methodology such as participant pools, task demands, experimental design, reading material, technology used, etc. Discrepancies can be large, even for similar teams running similar studies: Gould and colleagues (1987a) found that people read more slowly from screens than on paper, but discovered a short time later that screen readers could actually read just as quickly provided particular image quality thresholds (Gould et al., 1987b). Such findings illustrate just how nuanced the differences can be between screen reading and paper reading.

The conclusions by Gould's teams are not anomalous; Oborne and Holton (1988) claimed that "when all variables remain constant there is no difference in reading speed or comprehension between screen and paper" (p. 7), and that many other studies failed to hold all variables constant, thus introducing variability that could account for screen-paper performance differences found in those studies. Ackerman and Goldsmith (2011) agreed that comprehension differences are negligible between screen users and paper readers, but only when study time is fixed; the differences become significant in self-regulated study (unlimited study time), as paper readers generally perform better and choose to spend more time with the text. However, in terms of showing what they have learned, screen test takers generally complete assessments more quickly while scoring just as well as "paper and pencil" test takers (Bodmann & Robinson, 2004); other times, they actually score better (DeAngelis, 2000).

Differences between reading from screens and paper could be a product of the respective induced reading behaviors, with the implication being that the gap between the two media could be narrowed if people took the same attitudes toward reading in all of its forms. A survey by Liu (2005) found that "screen-based reading behavior is characterized by more time spent on browsing and scanning, keyword spotting, one-time reading, non-linear reading, and reading more selectively" (p. 700), acts that are not as common in print reading. Morineau et al. (2005) concluded similarly that people put forth less cognitive effort when using computerized environments. Perhaps as screens and other electronic media become more commonplace, people will read from them increasingly like they do paper texts (assuming that familiarity with the study medium is indeed a factor in reading behavior), thus narrowing any existing performance differences between users of the two types of study mediums.

1.2 Metacognition and learning

The previously-discussed advantages and disadvantages of screens and paper are most relevant when they are eventually related to the way people learn from the text. The most common way to assess learning and related factors is the comprehension test. The present study aimed to measure another important factor in learning in addition to comprehension: metacognition. Metacognition is important because it provides insight into the way people learn, can inform the way materials are designed, and is a key component to self-regulation.

According to Flavell (1976), metacognition is "one's knowledge concerning one's own cognitive processes or anything related to them." It consists of two types of skills (Brown, 1987): Knowledge of knowledge (i.e., figuring out a knowledge gap) and regulation of knowledge (i.e., devising a way to overcome the gap). Sometimes, the terms "metacognition" and "self-regulation" are used interchangeably, and they are often confused with each other, but they

actually refer to different processes. By most definitions, self-regulation is a general process that includes metacognition as a component. Zimmerman (1995) stated that self-regulation "involves more than metacognitive knowledge and skill, it involves an underlying sense of self-efficacy and personal agency and the motivational and behavioral processes to put these self-beliefs into effect" (p. 217). Furthermore, self-regulation often originates from the environment's stimulation of an individual's awareness, whereas metacognition "emphasizes learner development over learner-environment interactions" as the starting point for judgments and evaluations (Dinsmore, Alexander, & Loughlin, 2008, p. 393).

As described in the previous section, researchers have completed many studies regarding comprehension differences between screen readers and paper readers, but not as much has been done regarding differences in metacognition between the media. One of the findings from recent literature is as follows: When under time restrictions, only participants reading text from paper were able to become more efficient in their learning, scoring as well as they would have in conditions with no time pressure (Ackerman & Lauterman, 2012). Paper learners were also found to generally exhibit less overconfidence and comprehend better than screen learners, suggesting that "the primary differences between the two study media are not cognitive but rather metacognitive – less accurate prediction of performance (POP) and more erratic study-time regulation on screen than on paper" (Ackerman & Goldsmith, 2011, p. 18).

Research also exists on methods of enhancing the self-regulation processes of readers. Kauffman, Zhao, and Yang (2011) found that learners who experienced self-monitoring prompts in online text took better notes and scored higher on comprehension tests. In electronic environments, students given reflection prompts achieved higher self-regulation scores (van den Boom, Paas, van Merrienboer, & van Gog, 2004). Furthermore, Pressley (1995) hypothesized

that self-regulation benefits from relevant prompting can be embedded in extended learning periods (e.g. courses), not just one-time episodes. But prompts are not the only way of helping learners regulate themselves; self-questioning methods also enhanced various measures of self-regulated learning (Kramarski & Michalsky, 2010).

Of course, the ultimate goal for researchers is not just to find ways to improve self-regulation, but to convert improvement in self-regulation into improvement in comprehension and other more tangibly useful outcome measures. Some studies have indeed found that self-regulated learners tend to create good learning strategies that involve planning goals and selecting efficient approaches (Pintrich, 2000; Schraw, Crippen, & Hartley, 2006). Self-regulated learning also tends to produce increased performance for longer periods of time than non-regulated learning (Sitzmann, Bell, Kraiger, & Kanar, 2009).

The metacognitive prompts used in this study aided participants in self-regulation according to the concept of self-regulated learning competence (SRLC) presented by van den Boom, Paas, Merrienboer, and van Gog (2004), which is itself a refinement of Zimmerman's learning cycle phases (1998). The SRLC model outlines three stages in the study process: "starting," "performing," and "finishing." In order to properly self-reflect in each of these stages according to SRLC, learners should orient and plan when "starting," monitor and adjust while "performing," and assess and evaluate when "finishing." The thoughts elicited in the three stages are called "forethought," "intermediate thought," and "afterthought," respectively. In the reading portion of the study, participants in the "metacognitive prompts" conditions encountered a forethought prompt before reading each passage, an intermediate thought prompt during the passage, and an afterthought prompt at the conclusion of the passage. The prompts were designed to encourage the behaviors listed in the SRLC model, and the specific activities in the

prompts are based on a list of "constructive activities" proposed by Chi (2009). This study was also based on the work of Ackerman and Goldsmith (2011), who found that participants who read text from screens performed worse on two measures than those who read from paper: prediction of performance on a comprehension test (screen readers were generally overconfident and therefore exhibited lower metacognition) and comprehension (screen readers scored lower on comprehension tests).

This study manipulated the types of prompts used in texts (either metacognition-inducing SRLC prompts or non-metacognitive prompts) and the study medium (text presented either on screen or on paper) to find out their effects – whether main or interactive – on performance prediction accuracy, comprehension, study time, and reading behaviors. The design is therefore a 2x2 factorial with prompt type and study medium as the independent variables. In a sense, the study aimed to find – in addition to the effects of study medium – whether metacognitive prompts could compensate for some of the negative effects of screens (overconfidence, slower speed, lower comprehension to some extent, erratic reading behaviors, etc.).

Hypotheses prior to the study were made according to the discrepancy reduction model of learning (Butler & Winne, 1995; Dunlosky & Thiede, 1998), which describes learners setting a target level of learning and allocating study time according to a comparison of their subjective assessment of knowledge and the target level; when the target level is reached, people discontinue studying and move to the next step. To this end, participants were allowed unlimited time to read each passage, and the amount of study time participants devoted to each of the texts was recorded as a reflection of their metacognitive control decisions (i.e., when they believed that they were sufficiently ready to take the comprehension test).

1.3 Hypotheses

The previously-discussed research indicated that learners tend to demonstrate better metacognition when reading on paper and that metacognitive prompts are indeed helpful in promoting self-regulation, which is known to aid comprehension. Therefore, it was hypothesized that the participants in the "paper with metacognitive prompts" condition would score higher on the comprehension tests than those in all other conditions. Similarly, participants were expected to score lowest when reading from screens and using non-metacognitive prompts.

The expected differences in comprehension levels for participants in the other two conditions were more nuanced; the relative effects of the prompts and study medium had to be weighed against each other. It was hypothesized that because the participants in the present study were college-aged students, who presumably are accustomed to using technology in many facets of their lives, scores from participants in the "screen with metacognitive prompts" condition would be higher than those from participants in the "paper with non-metacognitive" condition. The implication of this hypothesis was that the increase in comprehension induced by the metacognitive prompts was expected to more than compensate for the predicted decreases in comprehension induced by the screen; after all, the literature yields mixed results in terms of study medium effects on comprehension. In simpler terms, prompt type was hypothesized to be a larger factor than study medium. It was also hypothesized that the metacognitive prompts would have greater comprehension effects for screen readers than paper readers because the participants reading screens would have more "room to grow" in terms of test scores. These hypotheses regarding comprehension are summarized qualitatively in Figure 1.

Comprehension
(score on post-test)

| Paper with metacognitive prompts | Screen with metacognitive prompts | Paper with non-metacognitive prompts | Screen with non-metacognitive prompts |
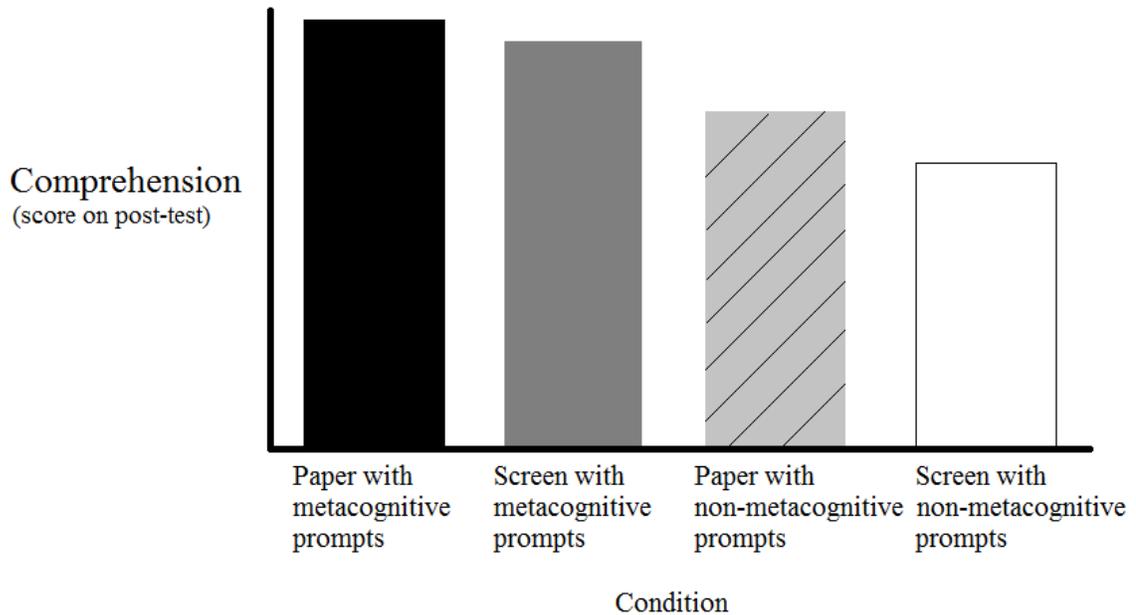
Condition

Figure 1. Hypothesized comprehension of participants by condition

In terms of accuracy in predicting performance, the hypotheses are relatively similar because self-regulation and comprehension are expected to be closely linked in that self-regulation aids comprehension. The "paper with metacognitive prompts" condition should produce the highest POP accuracy because of the participants' naturally-higher levels of metacognition on paper and the presence of metacognitive prompts to help them gauge their own understanding. Conversely, participants will predict least accurately in the "screen with non-metacognitive prompts" condition because using screens generally produces erratic reading (Liu, 2005) behavior and because the help of metacognitive prompts will not be available. POP accuracy for participants in the "screen with prompts" condition will likely be higher than those in the "paper with non-metacognitive prompts" conditions for the same reasons (i.e., effects of prompt type will be larger than those of study medium because of technology accustomedness) as explained in the previous paragraphs, and screen readers will show more improvement in POP when using metacognitive prompts because they generally have greater problems overestimating

their abilities than paper readers. The key difference between the hypotheses about comprehension and POP is driven by the fact that study medium has a larger effect on POP than on comprehension. Figure 2 qualitatively illustrates the hypotheses regarding POP.
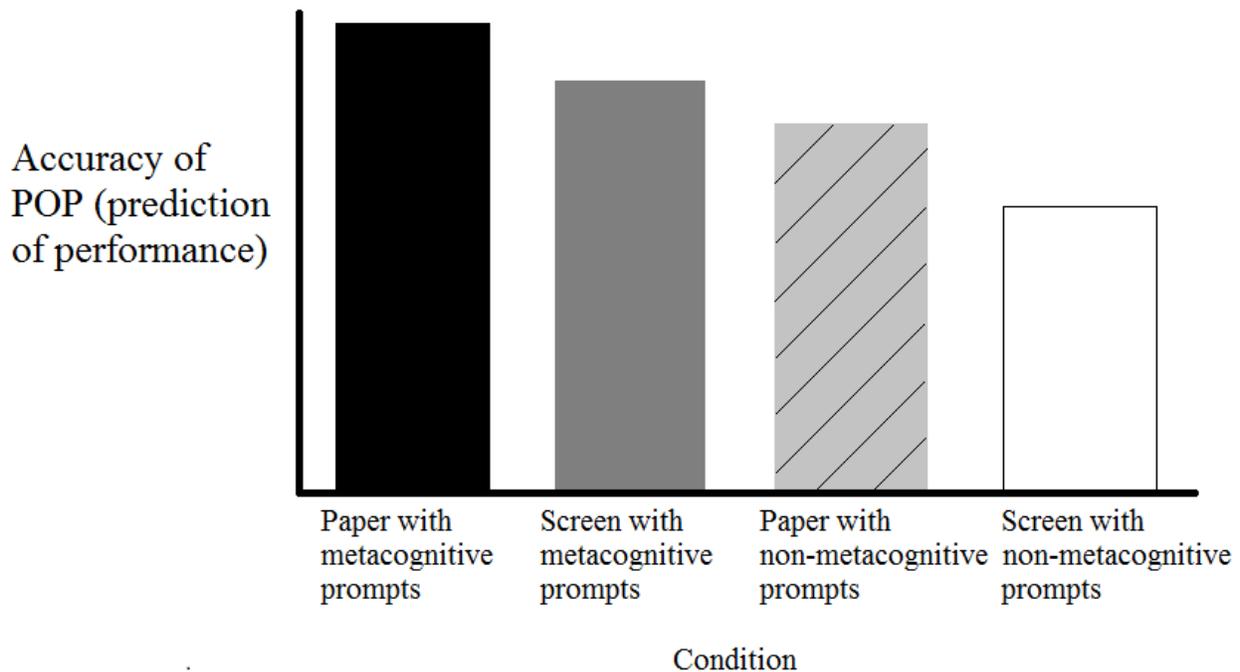
Figure 2. Hypothesized POP accuracy of participants by condition

Due to the tendencies of screen users to read in ways that resemble (or at least approach) "skimming", they are hypothesized to spend less time studying than those reading from paper. Using the same reasoning, screen users are also likely to produce lower-quality note-taking and prompt responses and engage in fewer non-required study behaviors. People reading from screens also tend to read more erratically – in selective or otherwise non-linear fashions (Liu, 2005) – and in ways not conducive to deep understanding, so they are not necessarily compensating for the time saved with greater rates of comprehension.

CHAPTER 2

METHOD

During the study, participants read an informed consent form, completed a demographics questionnaire, read text passages with prompts interspersed throughout, assessed their own comprehension of the texts, took tests on each of the passages, and completed a survey about their actions during reading.

2.1 Participants

Participants were 92 students (49 male, 43 female) from the Georgia Institute of Technology, all of whom received course credit for participation and were between 17 and 23 years in age ($M = 19.7$ years, $SD = 1.6$). The participants in each condition were equal in terms of pre-existing demographics such as grade-point average [$F(3, 59) = 0.91, p = 0.44$], SAT verbal scores [$F(3, 46) = 0.73, p = 0.54$], gender [$F(3, 88) = 0.87, p = 0.46$], and age [$F(3, 88) = 0.70, p = 0.55$].

2.2 Design

As stated previously, the design of this experiment was a 2x2 factorial with the independent variables being prompt type (metacognitive prompts or non-metacognitive prompts) and study medium (text presented either on paper or on a computer screen). Dependent measures included the amount of study time participants needed for each passage, comprehension test scores, participants' predictions of their test scores, the difference between predicted tests scores and actual test scores, self-reported study behaviors from the participants, and the quality of notes and prompt responses.

2.3 Materials

All students electing to participate in the study were given a demographics questionnaire to complete, which inquired about gender, age, college major, year in school, college grade point average, SAT scores, reading habits, screen usage, native language, and previous experience with the topics covered in the reading passages. The questionnaire was completed before the study began. Participants also received a stopwatch and instruction on how to use it to track the time they needed to read the passages.

During the study, all participants read three expository text passages, each of which was 500-1,000 words in length (the order of the texts was completely counter-balanced between participants and conditions to ensure that any order effects were distributed evenly across conditions). They read either from a standard computer monitor or from printed pages, depending on the randomly-assigned condition, and their texts contained either metacognitive prompts or non-metacognitive prompts interspersed throughout, to which the participants responded while reading. Text was presented, for both paper and screen conditions, in 12 type size and Times New Roman font with black letters appearing on a white background. However, minor extraneous discrepancies between the two conditions existed because screen condition participants had occasional contact with the computer interface, for which there was no analogous activity for paper condition participants. Furthermore, details about reader behavior were difficult to control, such as distance from text (about eighteen inches at the computer workstation but variable in paper conditions) or use of navigation mechanisms (scrolling via mouse or keyboard in screen conditions). The computers themselves had these specifications:

- Model: Dell Inspiron 570

- Monitor: Dell model number E172FPB, 17" LCD

- Resolution: 1280x1024

- Color: 32-bit

- Screen refresh rate: 60 Hz

While reading the passages, participants were required to respond to prompts interspersed throughout the text. If the participant was in one of the conditions with metacognitive prompts, then the prompts were designed according to SRLC specifications. As stated before, the specific activities in the prompts were based on Chi's "constructive activities" (2009), and these activities are shown in Table 1.

Table 1. SRLC stages and associated activities

SRLC stage (van den Boom, Paas, Merrienboer, and van Gog, 2004)	**Activity type (Chi, 2009)**
Forethought	"Organize own knowledge for coherence"
Intermediate thought	"Construct a concept map," "reflect, or self-monitor"
Afterthought	"Integrate new information with existing knowledge," "repair own faulty knowledge," "restructure own knowledge"

Examples of these metacognitive prompts are shown in Table 2 (every passage and the corresponding prompts are available for reference in the appendices):

Table 2. SRLC stages and examples of associated prompts

SRLC stage (van den Boom, Paas, Merrienboer, and van Gog, 2004)	**Example prompt**
Forethought	What do you know about the following topics (questionable treatment of animals, diseases, environmental impacts, and regulations) regarding animal factory farms? On the separate sheet designated for prompt responses and note-taking (in the Prompt 1 box), fill in – under the provided headings – any previous knowledge, experiences, or examples, you have related to these topics.
Intermediate thought	On the separate sheet designated for prompt responses and note-taking (in the Prompt 2 box), check your understanding so far by creating flow maps of these two processes: How animal meat becomes tender (paragraph 2) and how antibiotics get into our food (paragraph 3). Refer back to text and concept maps review sheet as needed
Afterthought	On the separate sheet designated for prompt responses and note-taking (in the Prompt 3 box), integrate your previous knowledge/experience with the information you learned from the article in a summary paragraph for each heading, and be sure to correct any misconceptions you might have had before reading the article (the box has the same headings as the one you filled in before reading the article).

For the non-metacognitive conditions, the prompts were designed to encourage the participants to reflect on the material, but not in a way that enhanced self-regulation of learning, per previous work by van den Boom, Paas, Merrienboer, & van Gog (2004). One example of this type of prompt is: "Do you think the rise of texting (and other short-form communication) is leading to a general decline in the ability of younger people to write properly and with nuance?"

To ensure that participants in the metacognitive conditions were correctly performing the activities in the prompts, examples of outlines and the appropriate concept maps were provided at the beginning of the study and available to the participants while reading the passages and responding to prompts. Those in the non-metacognitive conditions were not provided with that information because it did not apply to their prompts.

Non-metacognitive prompts were similarly placed throughout the texts for participants in those conditions. In all conditions, participants were required to respond to prompts by writing their answers on a sheet designated for prompt responses and note-taking. Prompt responses in all conditions were compared across the two study mediums as well as tested for comprehension test predictive power. Furthermore, as a manipulation check, the researchers graded metacognitive prompt responses on the accuracy of the concept maps and integration of old and new information, while non-metacognitive prompt responses were graded on completeness of answers (metacognitive participants scored 76% of available prompt-related points while non-metacognitive participants scored 92%, indicating that the prompts had produced relatively different behaviors). The notes taken by participants were also graded for thoroughness and inference-making (two-point scale: 0 = no notes taken, 1 = general ideas and/or scant notes, 2 = detailed ideas and/or plentiful notes).

After all participants finished reading a particular passage, they were provided with two example questions at the end of the text: one multiple-choice question for declarative fact knowledge and one short-response question for inference-making and/or information integration. Examples of these questions from a different article are shown below in Table 3 (for reference in the appendices, these questions are shown with each of their passages):

Table 3. Examples of post-passage example questions

Question type	**Example**

Multiple-choice	Which of the following, according to the author, does <u>not</u> change the "feeling" of a word?
	  a. Perceived eloquence
	  b. Speed with which word can be spoken
	  c. Number of letters in the word
	  d. Image conjured by the word
	  e. Number of syllables in the word

Short-response	The words "nosy," "inquisitive," and "curious," have similar meanings but different connotations. Rank these words in order of how favorable they are in terms of describing a person and defend your ranking with arguments that the author uses in the passage.

Participants were also asked to use those questions as guidelines to predict their performance on the ensuing comprehension test questions – this dependent measure is called "prediction of performance," or POP. This step was conducted after each passage and entirely on paper, regardless of the study medium used by the participant.

Comprehension tests took place after a participant finished a particular passage and completed the associated POP. The tests consisted of multiple-choice questions for the purposes of testing declarative fact memory and short answer questions to test inference-making and/or information integration. The materials for this step were also entirely on paper, regardless of the participant's study medium.

2.4 Procedure

All participants read the informed consent form and signed it before participating in the study. They were then given the demographics form to fill out and provided information to the extent that they are comfortable. After all of the initial paperwork was completed, the participants were permitted to begin the study.

The materials given to each participant were from one of the four conditions described earlier: "Paper with metacognitive prompts," "paper with non-metacognitive prompts," "screen with metacognitive prompts," or "screen with non-metacognitive prompts." In all conditions, participants received a sheet of paper to write down responses to their prompts. They also received electronic timers to track these two items:

- The time required to answer each prompt – time was used as a proxy for the amount of effort expended by the participant; therefore, the metacognitive and non-metacognitive prompts were compared on this measure to find out whether their differences were merely in the type of thinking prompted in the participant or whether the effort expended to answer the prompt was a potential confounding variable.

- The time from the start of a passage to the point at which they are ready for POP – this measure represents the time required by the participant to achieve the comprehension level he or she displays on the comprehension test.

Before reading began, participants were told that they had unlimited study time for each of the articles and the corresponding prompts and tests, although reading passages were waived on occasion when time constraints became an issue (experimental time slots were budgeted for two hours, although most participants finished before that time). The nature of the prompts was also explained at this time. During the reading periods, participants were allowed to take notes at any time, but performed these tasks in order, which were also written on the passages as instructions:

1. Start the timer for the passage
2. Respond to the first prompt (preceding the passage), then record time showing on timer

3. Read the passage until reaching the intermediate thought prompt in the middle of the passage, then record time showing on timer

4. Respond to second prompt, then record time showing on timer

5. Read the rest of the passage until reaching the third prompt (at end of passage), then record time showing on timer

6. Respond to third prompt, then record time showing on timer

7. Read example questions

8. Stop the timer when ready to proceed to the comprehension test, then record time showing on timer

When ready for the comprehension test, the participants alerted the researcher, who provided the POP materials at that time. After POP was complete, the researcher collected it and distributed the comprehension test, which the participant then completed. This step concluded the procedure for the first passage, and this procedure was repeated for each of the passages until all three were completed or time constraints were reached. At the end of the last comprehension test, the participants filled out a short survey about their actions during the experiment.

Each of the three comprehension tests were composed of five multiple-choice and three short-response questions much like the examples provided earlier. Multiple-choice questions were graded on an "all or nothing" basis while short responses could receive partial credit and were assessed by two graders and tested for inter-rater reliability (grades were reconciled through discussion when discrepancies were found). Below is an example of the grading scheme used for this short-response question: "If the author was in a car accident and narrowly avoided serious injury, would he use the word "happy" or "glad" to describe his emotions? Pick one of the words and defend your answer using information from the article."

- If picking "glad" (max = 1)

    - 0.25 | Picks "glad"

    - 0.25 | "Glad" is for relief/content

    - 0.25 | "Happy" is carefree (not applicable for this situation)

    - 0.25 | Describes why it's more of a relief than something to be carefree about

- If picking "happy" (max = 0.5)

    - 0.5, 0.25, or 0 | Make interpretation on convincingness

CHAPTER 3

RESULTS

The main objective of this study was to examine the effects of study medium (paper vs. screen) and prompt type (metacognitive or non-metacognitive) on various learning- and reading-related outcomes. The organization of this section and the next one will be centered on these manipulations and dependent measures.

3.1 Effects of study medium

<u>3.1.1 Reading comprehension</u>

The comprehension scores of paper readers ($M = 0.69$, $SD = 0.13$) and screen readers ($M = 0.69$, $SD = 0.17$), in terms of performance on multiple-choice questions, were not statistically different, $t(90) = 0.33$, $p = 0.74$ (Figure 3).
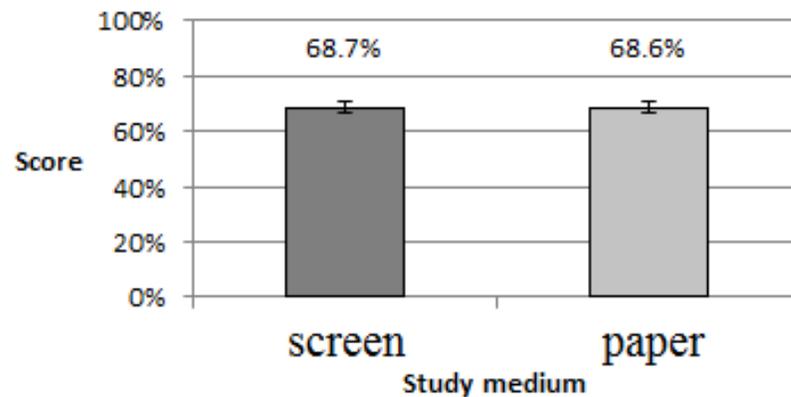
Figure 3. Multiple-choice comprehension scores compared across study mediums

A similar pattern of non-significant differences held for performance on short-answer questions as well, as seen as in Figure 4 below, $t(90) = 0.60$, $p = 0.55$ (Note: Short-answer responses were scored by two raters and analyzed for inter-rater reliability; the intra-class correlation coefficient of absolute agreement for reading comprehension scores was 0.96). These

results indicate that participants comprehended as much when reading text from screens as they did when reading text printed on paper.
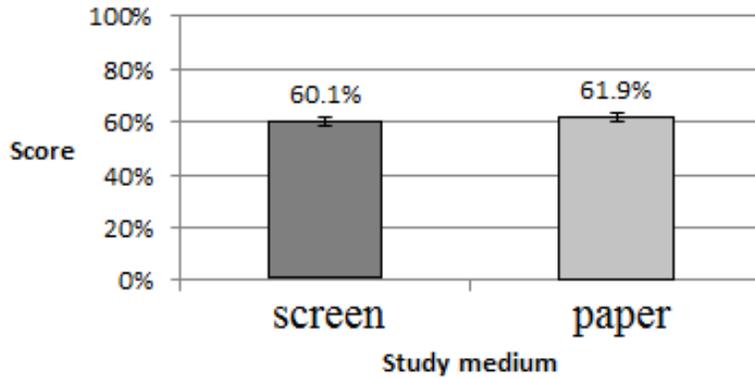
Figure 4. Short-answer comprehension scores compared across study mediums

To further investigate these results, ANCOVA analyses were performed to control for the effects of relevant covariates. The effects of study medium were still statistically non-significant when controlling for note-taking output, in terms of multiple-choice comprehension, $F(1, 89) = 0.16$, $p = 0.69$, and short-answer comprehension, $F(1, 89) = 0.15$, $p = 0.70$. Controlling for total study time also did not significantly change the effects of study medium for multiple-choice comprehension, $F(1, 89) = 0.22$, $p = 0.64$, and short-answer comprehension, $F(1, 89) = 0.35$, $p = 0.55$.

3.1.2 Prediction of performance (POP)

To calculate the accuracy of a participant's POP, the participant's actual test scores were subtracted from the participant's predicted scores to create a difference score for each test, a technique used by Wippich (1981). When the two study mediums were compared on this measure of prediction accuracy, the accuracy was not significantly different (in this case, lower means indicate better prediction accuracy); for multiple-choice questions, participants using screens ($M = 0.09$, $SD = 0.18$) over-estimated by about as many percentage points as those reading from paper ($M = 0.08$, $SD = 0.18$), $t(90) = 0.21$, $p = 0.84$; a similar pattern of statistical

non-significance was also observed in the short-answer realm, t (90) = 0.267, p = 0.79. The statistical differences between the absolute values of the difference scores (a technique used by Pressley et al.,1987) were also not statistically significant for multiple-choice questions, t (90) = 0.06, p = 0.95, and short-answer questions, t (90) = 0.35, p = 0.73.

A correlation analysis (predicted performance was correlated with actual performance; C. Hertzog, personal communication, March 5, 2015) further suggested that, for multiple-choice questions, neither those reading screens (r = 0.23, p = 0.12) or those reading paper (r = -0.22, p = 0.15) could predict future test performance with any statistically significant accuracy, with short-answer performance prediction again following suit on screen (r = 0.12, p = 0.44) and paper (r = -0.03, p = 0.85). However, a Fisher's R-to-Z transformation does reveal that screen participants were relatively better, in terms of these correlations, at predicting multiple-choice performance, z = 2.12, p = 0.03. Conversely, differences in the correlations for short-answer questions were not significant, z = 0.7, p = 0.48. In short, the results in total are a bit mixed: In the absolute sense, both groups are poor at predicting performance (as demonstrated by the correlation coefficients and their significance values) and overestimate on average by similar amounts (as demonstrated by the t-tests of the difference scores), but screen participants' predictions correlate better with actual performance on multiple-choice questions compared to predictions by paper participants (as demonstrated by the Fisher's R-to Z transformation).

3.1.3 Study behavior/preferences

Although the study medium manipulation appeared to have little effect on how readers comprehended text or assessed their own learning, some behavioral differences did emerge. The participants reading text on paper (M = 0.55, SD = 0.64) tended to take more notes while reading than those in the screen conditions (M = 0.26, SD = 0.49), t (90) = 2.12, p = 0.04 (Note:

Participants' notes were scored using a 2-point scheme awarding 0 points for not taking any notes, 1 point for general ideas and/or scarce notes, and 2 points for detailed ideas and/or abundant notes). These results can be seen in Figure 5 below.
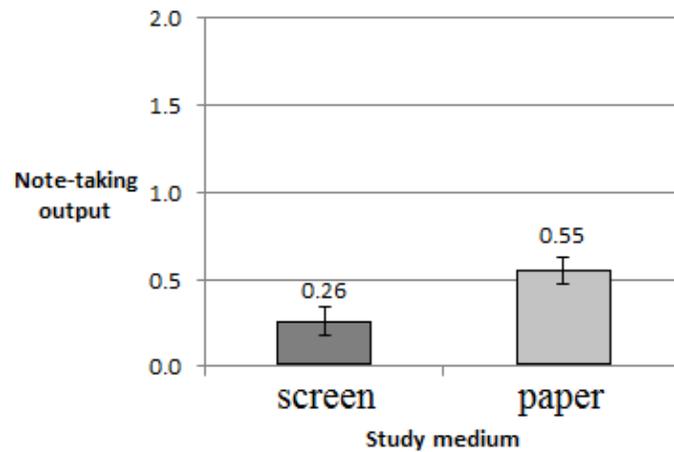
Figure 5. Note-taking output compared across study mediums

In addition, those in the paper conditions ($M = 19.48$ min., $SD = 6.58$) also spent more total time per article studying the text (including responding to prompts) than those reading from computer screens ($M = 16.43$ min., $SD = 4.41$), $t(90) = 1.96$, $p = 0.05$. These results can be seen in Figure 6 below.
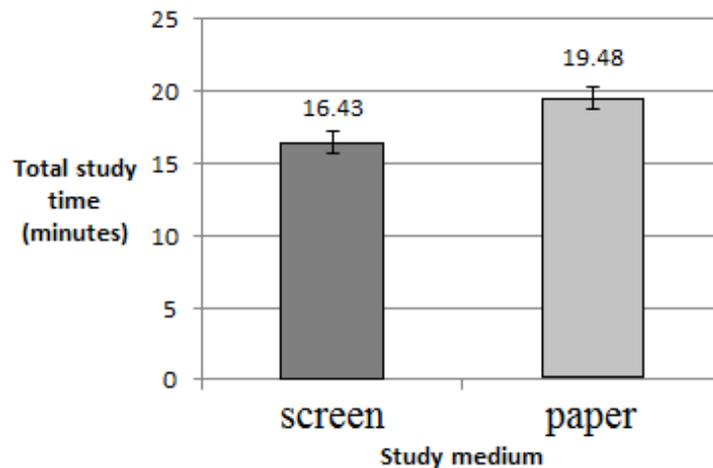
Figure 6. Total study time per article compared across study mediums

3.2 Effects of prompt type

3.2.1 Reading comprehension

Comprehension scores for those using the metacognitive prompts ($M = 0.71$, $SD = 0.14$) were not significantly higher than those using the non-metacognitive prompts ($M = 0.67$, $SD = 0.16$), in terms of performance on multiple-choice questions, $t (90) = 1.49$, $p = 0.14$, $d = 0.27$ (Figure 7). However, the effect size of this intervention is noteworthy and will be examined in the Discussion section.
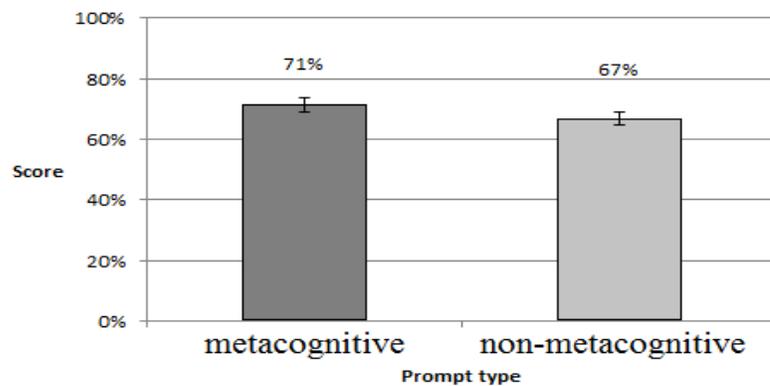
Figure 7. Multiple-choice comprehension scores compared across prompt type

Short-answer performance also failed to reveal any significant differences in comprehension, $t (90) = 0.33$, $p = 0.75$ (Figure 8).
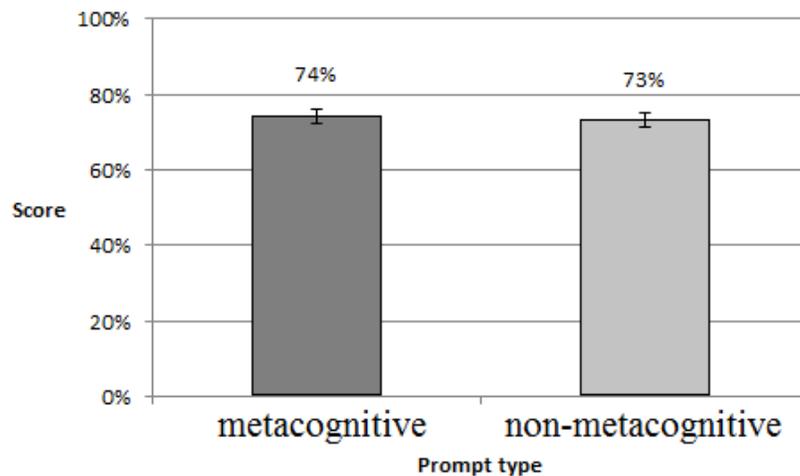
Figure 8. Short-answer comprehension scores compared across prompt type

According to these t-tests, the metacognitive prompts in text provided no statistically significant comprehension benefit to readers beyond what was already provided by non-metacognitive prompts. To further investigate the results, ANCOVA analyses were performed to examine whether statistical effects of prompt type might exist when controlling for relevant covariates. When controlling for note-taking output, prompt type still appeared to have no statistically-significant effects on multiple-choice scores, $F (1, 89) = 2.25$, $p = 0.14$, and short-answer scores, $F (1, 89) = 0.08$, $p = 0.78$. Controlling for total study time also left the effects of prompt type relatively unchanged for multiple-choice scores, $F (1, 89) = 2.06$, $p = 0.16$, and short-answer scores, $F (1, 89) = 0.16$, $p = 0.70$. On the other hand, after controlling for prompt response scores (a proxy measure for effort, as will be explained later), the marginal mean multiple-choice comprehension scores for participants using metacognitive prompts ($M = 0.76$, $SE = 0.03$) was significantly higher than those using non-metacognitive prompts ($M = 0.62$, $SE = 0.02$), $F (1, 89) = 12.21$, $p < 0.01$; the metacognitive participants also outscored the non-metacognitive participants in terms of short-answer scores, $F (1, 89) = 15.58$, $p < 0.01$.

3.2.2 Prediction of performance (POP)

Calculating POP accuracy was done the same way for prompt type as it was for study medium. For multiple-choice questions, the difference scores (measured as the difference between predicted performance and actual performance; Wippich, 1981) of participants using metacognitive prompts ($M = 0.08$, $SD = 0.17$) were found to be no different from the difference scores of those using non-metacognitive prompts ($M = 0.08$, $SD = 0.19$), $t (90) = 0.27$, $p = 0.84$; POP over-estimations for short-answer questions were also not significant, $t (90) = 0.27$, $p = 0.79$. An examination of the absolute values of the difference scores (Pressley et al., 1987) also

revealed non-significant differences in POP accuracy for multiple-choice questions, t (90) =

0.80, p = 0.42, and short-answer questions, t (90) = 0.68, p = 0.50.

Correlations between POP and actual performance (C. Hertzog, personal communication,

March 5, 2015) also revealed that participants were generally not able to reliably predict their

performance on multiple-choice questions, regardless of whether they used metacognitive

prompts (r = 0.22 , p = 0.15) or non-metacognitive prompts (r = -0.17 , p = 0.25). However, the

difference between those two correlation coefficients is statistically significant, indicating that

participants using the metacognitive prompts were, according to Fisher's R-to-Z, more accurate

in predicting performance than those using the non-metacognitive prompts, z = 1.83, p = 0.03.

Predictions of performance on short-answer questions were generally unreliable for

metacognitive prompts (r = 0.07, p = 0.64) and non-metacognitive prompts (r = 0.02, p = 0.90),

and the groups were not significantly different in this regard. Again, these results are a bit mixed:

According to difference scores, neither prompt type appeared to elicit particularly accurate

predictions for either question type. However, according to Fisher's R-to-Z, participants were

able to predict multiple-choice performance significantly more accurately when using the

metacognitive prompts.

3.2.3 Study behavior

Like the study medium manipulation, prompt type did appear to have some effects on

participants' behaviors during the study. For example, as was mentioned previously, participants

exposed to the metacognitive prompts reported re-reading text more often than those using non-

metacognitive prompts. That difference could be explained by the fact that those in

metacognitive conditions rated the act of referring back to the text and finding previously-read

text as significantly easier (M = 4.36, SD = 0.74) than those in the non-metacognitive conditions

($M = 3.81$, $SD = 0.88$), t (90) $= 3.22$, $p < 0.01$. Alternatively, participants in the non-

metacognitive conditions reported mentally summarizing the material along the way more

frequently ($M = 3.98$, $SD = 0.92$) than those in the metacognitive conditions ($M = 3.31$, $SD =$

1.13), t (90) $= 3.12$, $p < 0.01$. These data reveal that the key difference between the prompts is in

the nature of the behavior induced in the participants; the participants in the metacognitive

conditions could more easily refer back to and re-read text and therefore preferred to do that,

while those in the non-metacognitive conditions generally chose to summarize the material while

reading. Figure 9 displays the frequency-related results graphically.
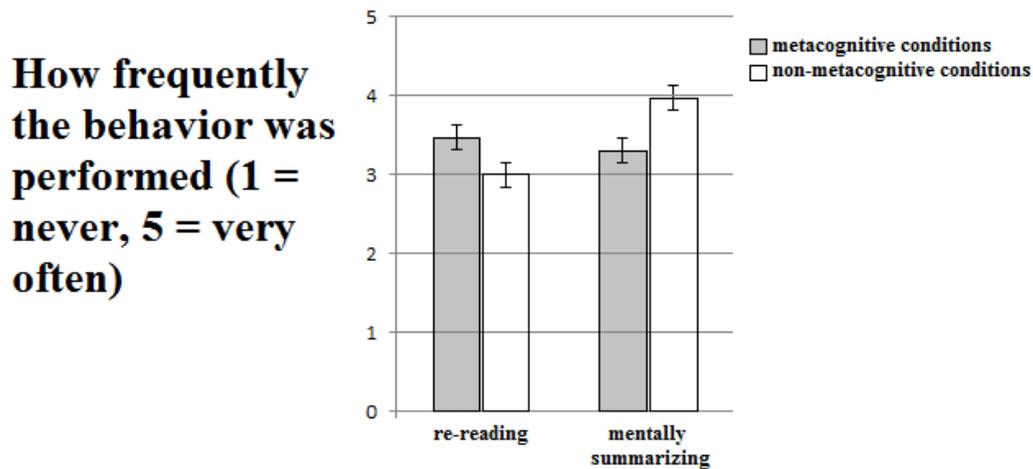
Figure 9. Frequency of re-reading and mentally summarizing by prompt type

CHAPTER 4

DISCUSSION

4.1 Effects of study medium

4.1.1 Reading comprehension

Although it was hypothesized that those reading from a screen would comprehend less than those reading from paper, the results did not support that hypothesis. The results are in line, however, with recent trends in the "paper vs. screen" literature, which demonstrate that differences in how people read on screen and on paper are diminishing as technology becomes more commonplace. For example, Noyes and Garland (2008) concluded in their meta-analysis that although equivalence between the two mediums is inherently difficult to achieve, greater equivalence is happening more now than in past generations. This trend toward equivalence is not necessarily a surprising one; after all, technology has become a much larger presence in the lives of most people, enabling them to become almost as familiar with computer screens as they are with paper. For the particular population examined in this study (college students), equivalence is even more likely because computers likely played significant roles in their formative years.

Research has shown that younger people are indeed more comfortable with computers and generally have more feelings of control than older people (Czaja & Sharit, 1998). In fact, students are more likely to use e-books than faculty and staff (Anuradha & Usha, 2006). In the present study, that comfort level manifested itself in the fact that those reading text from computer screens comprehended just as much from paper. The effects of technology familiarity also seemed to extend into areas that could indirectly affect comprehension, such as mental demand; as measured by NASA TLX responses, participants reported statistically-similar

28

amounts of mental demand in the screen conditions ($M = 0.55$, $SD = 0.18$) as in the paper

conditions ($M = 0.50$, $SD = 0.19$), t (89) = 1.27, $p > 0.05$. The deleterious effects of extraneous

cognitive load are well-documented (e.g. Paas, Renkl, & Sweller, 2003), and it was thought that

perhaps computer screens would produce some extraneous load that could decrease

comprehension, but the difference in mental demand between the conditions appears to be

negligible.

Some of the directions of future work related to these findings are fairly easy to envision.

For example, a study similar to this one can be carried out using older populations to determine

whether their relative unfamiliarity with technology makes a difference in reading

comprehension. Another potential direction is to vary the genres of texts read by the participants.

In the present study, participants read human interest articles that were intended for academic

purposes and the results demonstrated that study medium had no significant effect on

comprehension, but perhaps other genres of texts could yield different results. Previous research

has found that people usually prefer paper when deep processing is needed (Wu & Chen, 2011).

Short stories or popular culture articles might be examples of text genres that elicit shallower

processing than academic articles, and this processing could interact with the study medium

used.

4.1.2 Prediction of performance

In general, paper participants did not demonstrably predict performance more accurately

than screen participants, and neither group was accurate in the absolute sense, contrary to the

expected results. In fact, the only statistically significant finding in this regard showed that

screen participants were actually relatively better than the paper participants in predicting

multiple-choice performance, a finding revealed when analyzing correlations between POP and

actual performance. These largely null results could be attributed to the fact that there was no significant difference in the difficulty of reading on screen and paper, as reported by the participants on a NASA TLX post-experiment survey. More specifically, the researcher originally hypothesized that participants using screens would have problems with overconfidence because of the "hard-easy" effect (e.g. Lichtenstein, Fischhoff, & Phillips, 1982), the tendency for people to be overconfident when studying difficult materials; a possible explanation for this effect is that working harder to understand material lends itself to feelings of more accomplishment, and therefore, feelings of superior understanding. However, as alluded to earlier, the difficulties inherent in reading from screens are diminishing still, particularly for those who have grown up with technology for most of their lives, leading to relative paper-screen equivalence in POP (or even slight advantages when reading screens) for this study and no screen-induced overconfidence.

Hard-easy effects due to study medium have been found in previous research (e.g. Ackerman & Goldsmith, 2011). Therefore, to further investigate the effects of study medium on POP, future studies could vary the image qualities of screens to induce hard-easy effects. Although the hard-easy effect did not seem to be at work in the present study, only one particular computer setup was used, hardly a representative sample of all computers and screens. Perhaps with screens of lower image quality (e.g. those used in developing countries), people would be working harder to read the text and thus become overconfident about their understanding of the material, leading to worse POP. Image quality of text on paper is also a possible future manipulation, although perhaps of less interest because printing quality is not usually a limiting factor in most learning environments.

The findings related to note-taking and study time suggest that A) when using paper to read text, participants were more deliberate and maybe more conscientious than when using screens (as noted previously, a participant's study time can be interpreted as a metacognitive control decision that he or she is ready for the test), and/or B) the participants' comfort level with technology has reached a point in which they are possibly more efficient in learning information on screens. After all, screen participants achieved comparable comprehension to paper participants while spending less time and taking fewer notes. According to research by Morineau et al. (2005), explanation A would not be surprising because people are inclined to work harder to understand text when it is presented on paper. The reverse also seems to be true – when people want to read in greater depth, they prefer to print out the text rather than read it on a screen (Wu & Chen, 2011). Jabr (2013) added that electronic text is more ephemeral in nature than printed text, possibly leading readers to feel that words in print are to be lent more credence than words on a screen. Future work could perhaps employ eye-tracking to investigate this conscientiousness in more detail; for example, when people read text on paper, what types of reading behaviors are done with the extra time?

Some of the previously-cited literature about the recent trend toward paper-screen equivalence does provide some merit to explanation B, which is not necessarily mutually exclusive with explanation A. That is, as people read more often on screens, whether by choice or necessity, they have adapted to modern circumstances and overcome some of the inherent screen-related difficulties whose existence previous research has demonstrated. However, there is one area in which screens have not "caught up" to paper yet: preference. Data from the post-experiment surveys show that the subjective experience of reading on screens is, although

familiar, still not preferred to reading paper. In the surveys administered after participants finished the study, a higher percentage of those in the screen conditions ($M = 46\%$, $SD = 50\%$) reported that they would have preferred using the other study medium to read the articles than those in the paper conditions ($M = 9\%$, $SD = 29\%$), $t(90) = 4.24$, $p < 0.001$. That is, screen participants wanted to switch to paper significantly more often than paper readers wanted to switch to screens. If the screen experience is less pleasant on a subjective level, it makes sense that those subject to the screen would want to end the experience more quickly than those who received paper.

An interesting direction for future studies on reading behavior involves the tablet, a relatively new form of technology that, in many ways, bridges the gap between screens and paper. The tablet presents text on a screen, but readers can use it to replicate some of the experiences of reading from paper such as associating information with its location on a particular page, flipping discrete pages, or holding the entire text in one's hands. The points about locating information and flipping discrete pages are important for comprehension: People often recall information by remembering where they saw it on a page (Rothkopf, 1971), which is more difficult to do when reading text that has to be scrolled through in a continuous stream. Scrolling is an interesting area of study because according to Dyson and Haselgrove (2000), how a person does it through text can be associated with his or her comprehension of the text; according to their research, the more time someone spends in between scrolling movements, the higher his or her comprehension tends to be. Scrolling text is also likely more mentally taxing than flipping pages because scrolling requires some focus on how the text is moving (Wastlund, 2007). However, tablets still fall short in some areas when compared to paper. Gerlach and Buxmann (2011) have proposed that digital books produce "haptic dissonance" because they do

not feel the way books are expected to feel, and this dissonance can cause discomfort in readers (although this dissonance is likely being reduced more by younger generations and as technology becomes more ubiquitous); for example, a digital book weighs the same regardless of the length of the book being read, while the weight of a printed book tends to have some relation to its length.

The tablet, as an intermediate medium, could facilitate future research that can be more discerning about the particular characteristics of screens and paper that produce certain outcomes. That is, by using tablets, researchers would be better able to account for variation in comprehension, study behavior, etc. by pinpointing the characteristics that create said variation. For example, the present study found that participants spent more time studying when reading paper than when reading from a screen; if participants also spend more time when using tablets than when using screens, then it can be inferred that screens themselves are not a deterrent to spending more time, and that perhaps discrete pagination is a feature that promotes increased study time.

4.2 Effects of prompt type

4.2.1 Reading comprehension

In terms of statistical significance, the results in this study did not support the original hypothesis, which stated that the metacognitive prompts would help readers improve comprehension. However, according to an effect size analysis, the participants using metacognitive prompts did outscore their non-metacognitive counterparts by 0.27 standard deviations for multiple-choice questions. To put that number in context, a meta-analysis of educational interventions by Lipsey et al. (2012) revealed that instructional format interventions (broad pedagogical strategies) yield median effect sizes of 0.13 standard deviations, while

teaching technique interventions (specific pedagogical strategies) generally yield 0.27 standard deviations of improvement. Therefore, even though the improvement in performance due to metacognitive prompts was not statistically significant, the size of the improvement is as large or even larger than that of many educational interventions in existing literature.

As for why the performance gap between the two groups was not significant in the statistical sense (aside from concerns related to sample size), one of the factors could be mental demand. On a Likert post-experiment survey question (1 = never, 5 = often), participants who read the text with metacognitive prompts ($M = 3.47$, $SD = 0.92$) reported re-reading text more often than those who were using the non-metacognitive prompts ($M = 3.00$, $SD = 1.00$), t (90) = 2.33, $p = 0.02$. According to a correlation analysis, this extra re-reading is associated with higher mental demand, as measured by NASA TLX ($r = 0.33$, $p < 0.01$). Therefore, the possibility exists that the metacognitive prompts did in fact induce metacognitive processes and contribute to better comprehension in participants, but also that the benefits were negated to some extent by the additional mental workload required to interact with the prompts.

Another possible explanation for the non-significance is related to the genre of text read in this study. Text genre was mentioned as a direction for future investigation of study mediums, but it could also have affected the results related to prompt type. More specifically, the benefits provided by the metacognitive prompts might not be as evident in particular genres of text. In this study, participants read human interest articles, in which later material does not inherently build as much on previous material; however, technical material like math and science does tend to build on earlier concepts. Perhaps metacognitive processes, such as a reader assessing how well he or she has learned earlier material, are more useful to readers when learning material that requires a cumulative understanding of earlier concepts.

Despite the statistical non-significance of these score differences between prompt types, the size of the manipulation effect does offer promise for future interventions related to metacognition and metacognitive prompts. After all, much research has revealed the benefits of metacognitive processes to learners, whether in note-taking (Kauffman, Zhao, & Yang, 2011), goal-setting (Pintrich, 2000; Schraw, Crippen, & Hartley, 2006), or long-term learning (Pressley, 1995), among other benefits. Therefore, further investigation of metacognition, and prompt-related interventions in particular, is warranted. Along those lines, future research in the area of metacognitive prompts could examine different manipulations of text genres. For example, material that tends to build later concepts from earlier concepts is expected to be better suited for usage of metacognitive prompts than material that does not necessarily build. Another approach to increase the effectiveness of the metacognitive prompts is to decrease the mental workload that was shown to be associated with using the prompts. Because participants did not necessarily comprehend more after incurring the increased workload of the metacognitive prompts, the additional load can be considered extraneous and possibly disruptive to learning (Paas, Renkl, & Sweller, 2003). For maximum benefit, prompts that encourage readers to assess their own learning should be designed to cause as little disruption as possible.

A natural limitation of this study is that it cannot examine every prompt-related method of encouraging metacognitive processes in readers; therefore, the results from this study should not be interpreted necessarily as a general statement on the usefulness of prompts in eliciting metacognition. For example, instead of the SRLC framework, the influential metacognitive model proposed by Nelson and Narens (1994) could be used to inform the creation of metacognitive prompts. These prompts would incorporate the two levels of processing that Nelson and Narens (1994) believe are necessary for metacognition: object-level (aiding specific

cognitive functions such as phonological coding and object recognition) and meta-level (aiding

evaluation of object-level activations and providing feedback). Of course, metacognition can also

be elicited using mechanisms other than in-text prompts, and these mechanisms are worth

exploring as well. For example, readers could be told, before reading an article, to employ a

particular learning strategy throughout the duration of the article, as opposed to being prompted

at various points in the article.

<u>4.2.2 Prediction of performance</u>

The researchers hypothesized that metacognitive prompts would help participants gauge

their own comprehension, and by one measure, the participants using them were better at

predicting performance on multiple-choice questions than participants using non-metacognitive

prompts. However, the results in general show that neither prompt type elicited particularly

accurate predictions, even if a difference for multiple-choice POP exists. Some explanations for

the results revolve around limitations of the study itself. For example, due to time concerns, only

two example questions (one multiple-choice, one short-answer) were provided at the ends of

articles for participants to use in predicting test performance, which might not have been enough

information for the participants to accurately predict performance, regardless of the type of

prompt received. Perhaps the POP effects of the metacognitive prompts would have been clearer

if sufficient example questions were included. Furthermore, the variation in effort on prompt

responses (which will be covered more in following sections) was noticeable, at least on an

anecdotal level; therefore, the possibility exists that in this study, the amount of metacognition

performed by participants was largely a function of prompt response effort rather than prompt

type, pointing perhaps to participants' variable motivation levels as a limitation in this study.

This explanation is consistent with research by Aleven and Koedinger (2000), which found that students' spontaneous usage of provided help is generally intermittent.

A third limitation is that participants were tested only immediately after reading each article, and not tested on later dates. No tests were done on later dates because it was thought that participants would have little motivation to remember the information past the experiment day, and such a situation would bear little resemblance to most learning environments in which people have at least some sort of intrinsic or external motivation to remember information for extended periods of time. Previous research has shown that self-regulated learning exhibits a permanence that non-regulated learning does not (Sitzmann, Bell, Kraiger, & Kanar, 2009). Therefore, it is plausible that the effects of better metacognition, insofar as the metacognitive prompts can induce it, are more evident in tests on later dates, whether the effects are on comprehension or POP.

Providing more/better example questions and long-term retention tests are relatively easy ways to deal with some of the aforementioned limitations. Making these corrections could reveal that metacognitive prompts are in fact useful in eliciting accurate POP, and that usefulness could manifest itself in statistical differences on multiple measures of prediction accuracy, as opposed to just one measure for one question type in the present study. However, other future work might explore the motivation/effort aspect of the study, which is somewhat more difficult to do. Whether the dependent measure is comprehension, POP, or participant behaviors, laboratory experiments of this kind possess an inherent difficulty in requiring participants to replicate how they would read for school and/or work. That is, participants might not do what they would normally do while reading and working, which decreases the generalizability of any findings from the study. For POP specifically, participants did not have much of a reason to strive for

accuracy in their predictions; in contrast, while answering comprehension test questions, participants are more likely to have motivations for doing well (e.g. learning something new, social comparison). Furthermore, the possibility exists that the metacognitive prompts were generally not as engaging as the non-metacognitive prompts. The non-metacognitive prompts usually asked participants about their opinions regarding issues, which might have been more intrinsically interesting and helped keep the participants engaged, even if the prompts were only tangentially relevant to the topic at hand; on the other hand, the metacognitive prompts could have felt like "extra" work. Both of these issues likely decreased how often participants assessed their own learning and how much effort they put into doing those assessments. Any future studies in POP must incentivize participants for accurate predictions so that lack of effort does not become a possible dominating factor.

4.2.3 Study behavior

The data show that re-reading was performed more often in metacognitive conditions, while mentally summarizing (while moving through the text) was performed more often in non-metacognitive conditions. A probable explanation for these behaviors is that the metacognitive prompts caused participants to assess their own understanding of the material at various points in the article, and in turn the participants referred back to the text accordingly. The re-readings likely helped participants to remember the location of previously-read ideas on the pages (Rothkopf, 1971), which explains why the metacognitive participants found referring back to be relatively easy. In contrast, participants in the non-metacognitive conditions were more likely to summarize the material as they read because they did not have the natural checkpoints that their metacognitive counterparts had. Naturally, because the non-metacognitive participants were summarizing more as they read, they likely did not feel the need to re-read as much.

38

Although the aforementioned data are somewhat mixed in terms of demonstrating the effectiveness of metacognitive prompts, some data from this study showed that the metacognitive prompts could be more effective given the right situation. According to a correlation analysis, metacognitive participants' scores on prompt responses are significantly correlated with their comprehension test scores for both multiple-choice questions ($r = 0.42$, $p = 0.004$) and short-answer questions ($r = 0.67$, $p < 0.001$); non-metacognitive participants saw positive correlations as well (average r-value $= 0.36$, average p-value $= 0.023$), but the correlations were not as strong. That is, the higher a metacognitive participant's prompt response quality was, the higher he or she scored on the comprehension tests. This finding is corroborated by the fact that prompt type did appear to have a significant effect on comprehension once prompt response quality was controlled for as a covariate. Given that both types of prompts were designed to be relatively straightforward, participants' scores on them can be interpreted as a proxy measure of effort. Therefore, it can be inferred that effort on a metacognitive prompt is more closely associated with higher comprehension than effort on a non-metacognitive prompt. Unfortunately, finding pre-existing individual differences that correlate well with prompt response quality has proven to be difficult; for example, neither a participant's SAT Verbal score ($r = 0.02$, $p = 0.91$) or college grade-point average ($r = -0.02$, $p = 0.89$) predicts particularly accurately whether a participant will put forth the effort to answer prompts well.

These correlations and ANCOVA analyses further solidify the idea that the metacognitive prompts, if used to their full potential (i.e., participant puts forth requisite effort in responding and paying attention to the content), can be useful. Given the ample manipulation effect size and the fact that participants' efforts on metacognitive prompts are more strongly linked to better comprehension than efforts on non-metacognitive prompts, it stands to reason

39

that the induced metacognition indeed yielded some positive results, as long as the participant was willing to put forth the effort. Research in the future could examine how to create metacognitive prompts that would encourage readers to actually put forth that effort. As mentioned previously, effort on the prompts (as measured by prompt response scores and anecdotally) varied widely among participants and likely contributed to the insignificant comprehension differences between participants in the metacognitive and non-metacognitive conditions. One way to attack the problem is to create "motivation profiles" (e.g. obtaining information about a participant's belief in own ability, expected success level, perceived difficulty and importance of subject, whether ability in subject is innate), which can then be used to tailor prompts and other learning aids for learners (Beal, Qu, & Lee, 2006). Motivating learners is crucial because some research has shown that motivation is a key factor in learning achievement (Byrnes, 2003).

Given that the prompts seem to elicit different reading behaviors, other future work in this area could focus on examining how those different behaviors affect performance. Perhaps behaviors such as re-reading and mentally summarizing can be examined and compared further with eye-tracking technology, which would hopefully allow researchers to examine particular traits of the behaviors that tend to produce improved performance. Prompts could then be designed to facilitate those behaviors that have been found to improve performance. Eye-tracking also eliminates the subjective nature of the data. For instance, in a self-report situation, a participant might report re-reading more often because he or she knows that it's the "right" thing to do in this study; eye-tracking could produce data that are more representative of actual learner behaviors.

CHAPTER 5

CONCLUSION

Ackerman and Goldsmith (2011) have posited that comprehension differences between learners using paper and those reading from screens are driven partly by metacognition, citing decreased POP accuracy among other measures; more specifically, they found that screens induce less metacognition in readers than paper. One of the main purposes of this study was to examine whether metacognitive prompts could perhaps compensate for this shortcoming of computer screens.

Existing literature has demonstrated that computers are "catching up" to paper in many reading performance measures. The results in the present study followed that trend, indicating that participants largely comprehended and predicted performance equally well whether reading from screens or paper, and by one measure actually predicted more accurately on screens. Participants in the present study also took more notes and spent more time studying when using paper, but this extra work failed to generate better performance, leading to the conclusion that screen participants performed more efficiently by achieving comparable comprehension and prediction accuracy with less work. Given that people still subjectively prefer paper over screens, in this study and elsewhere, Andre and Wickens (1995) might have been right that people sometimes do not want what is best for them.

As for the prompt type manipulation, what is best for people could be metacognitive prompts, although additional study is necessary. The effect size for metacognitive prompts on multiple-choice questions was comparable to the effect sizes found for many educational interventions in existing literature, signifying that metacognitive prompts produced better multiple-choice performance. Furthermore, controlling for prompt response scores revealed that

metacognitive prompts actually produced better performance on both multiple-choice and short-answer questions, provided that participants put forth the effort to answer the prompts. However, the statistical non-significance of the raw performance differences still leaves room for further investigation. The data were more definitive concerning some of the reading behaviors elicited by the different prompt types: Metacognitive participants tended to refer back to and re-read text, while non-metacognitive participants (whose prompts did not regularly check their understanding) were more likely to summarize as they read.

The results from this study can be added to ongoing discussions about the effects of technology in the classroom and how to design learning materials. These discussions, as they are related to metacognition, are important. According to Zimmerman and Martinez-Pons (1986), usage of self-regulation processes is a very accurate predictor of standardized test achievement.

The prevalence of technology-centered learning is has been on the rise for some time – in 2011, almost one-third of US college students had taken at least one online course (Online Learning Consortium, 2012). Research regarding the effects of various text delivery mechanisms is therefore important, as is research concerning instructional design that can help students regulate their own learning when a teacher is not present. Hopefully, this study provides some insight into both of these realms.

APPENDIX A: PREDICTION OF PERFORMANCE (POP) FORM

Topic of article: _____

1. What percentage of multiple-choice questions on the comprehension test do you believe you will be able to answer correctly? Use the difficulty of the example multiple-choice question (at the end of the article that you just read) as a reference.

Write your response here: _____ %

2. What percentage of short answer questions on the comprehension test do you believe you will be able to answer correctly? Use the difficulty of the example short answer question (at the end of the article that you just read) as a reference.

Write your response here: _____ %

You will now proceed to the comprehension test for the article you just read. Please alert the research assistant so you can be given the test.

Participant number: ____

Demographics questionnaire

Gender (circle one):   Male   Female

Age: ____

Major: _____

Year in college: First   Second   Third   Fourth   Fifth   Other: _____

College GPA (if you remember):   ____ / 4.0

SAT scores (if you remember):   Math ____   Verbal ____

How many hours per day do you think you spend looking at a screen of some kind (e.g. mobile phone, tablet, computer, television)?

None   1-2 hours   3-5 hours   6-10 hours   11+ hours

Is English your first language?

                    Yes   No

Have you studied post-traumatic stress disorder (PTSD) in school before, or have you or anyone close to you ever been diagnosed with PTSD?

                    Yes   No

How much knowledge do you have about animal factory farm practices?

      None   A little   Some knowledge   A lot

How much time do you spend reading per week (including books, magazines, online articles, newspapers; not including shorter items like tweets and Facebook statuses)?

None   1-2 hours   3-5 hours   6-10 hours   11+ hours

When speaking or writing, how conscientious/careful are you in terms of choosing the "perfect" word that is most appropriate for the context?

Circle one: 0   1   2   3   4   5   6   7   8   9   10
          Not at all   Obsessed

APPENDIX C: CONCEPT REPRESENTATIONS REVIEW SHEET

During this study, you will be making use of flow maps, double-bubble maps, and bar graphs. These concept representations are likely familiar to you, but you should read these descriptions/examples just to make sure:

Flow map: A flow map is used to outline the steps of a process, with each step being written in its own box. Below is an example of a flow map of the scientific method.

http://www.kleinisd.net/users/0082/docs/32-pd.flow.pdf
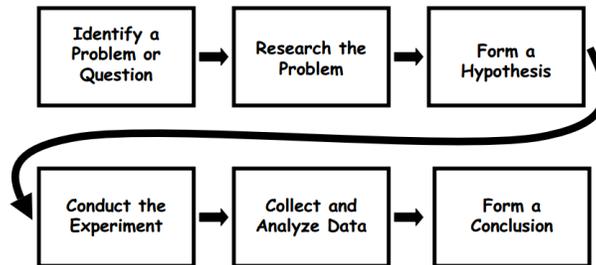
Double-bubble map: A double-bubble map is used to outline similarities and differences between two things. Similarities are written in the shared bubbles in the middle, while differences are written on either side according to the thing they describe. Differences associated with each other are often written in analogous locations on either side of the thing (e.g. in the example below, the functions' quadrants are directly above the bubbles for both of the functions).

http://gchstlc.pbworks.com/w/page/7076364/Double%20Bubble%20Maps
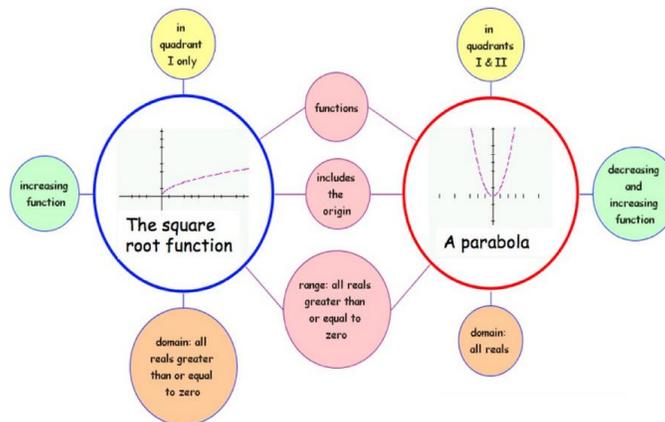
Bar graph: A bar graph is used to visually represent quantities via lengths of rectangular bars. Below is an example of a bar graph.

http://www.eduplace.com/math/mw/background/5/06a/te_5_06a_graphs_ideas.html

APPENDIX D: POST-EXPERIMENT SURVEY

<u>Post-experiment survey</u>

Circle one choice for each statement below.

1. The speed I read at was _____ my normal reading speed:

1	2	3	4	5
much slower than		about the same as		much faster than

2. I re-read some parts of articles to ensure I understood concepts or to refresh my memory.

1	2	3	4	5
never	rarely	occasionally	sometimes	often

3. My level of concentration was _____ my usual concentration when reading for school.

1	2	3	4	5
much lower than		about the same as		much higher than

4. I skimmed roughly _____ of the text in the articles.

Circle one: 0%   10%   20%   30%   40%   50%   60%   70%   80%   90%   100%

5. I was able to easily refer back to the article to find text I read earlier.

1	2	3	4	5
fully disagree		neutral		fully agree

6. Read the question description, and then write an "X" on the gradation line that corresponds to your answer.

MENTAL DEMAND   How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

MENTAL DEMAND

Low   High

7. I set comprehension goals before starting each article in this study: YES   NO

8. How well do you think you performed on the post-tests (points earned divided by points possible)?

Circle one: 0%   10%   20%   30%   40%   50%   60%   70%   80%   90%   100%

9. The articles took _____ expected to read (given that they were about 800-1000 words in length).

1	2	3	4	5
much less time than		as long as		much longer than

10. I was able to relate information in the articles to my previous knowledge/experience.

1	2	3	4	5
fully disagree		neutral		fully agree

11. I self-questioned my understanding of the material while I was reading.

1	2	3	4	5
never	rarely	occasionally	sometimes	often

12. I mentally summarized the material while reading.

1	2	3	4	5
never	rarely	occasionally	sometimes	often

13. I felt prepared to read the articles after completing the first prompt (before reading).

1	2	3	4	5
fully disagree		neutral		fully agree

14. I made changes in my concentration and/or strategies (whenever needed) while reading the articles.

1	2	3	4	5
fully disagree		neutral		fully agree

15. Read the question description, and then write an "X" on the gradation line that corresponds to your answer.

PERFORMANCE   How successful do you think you were in accomplishing the goals   PERFORMANCE
of the task set by the experimenter (or yourself)? How satisfied
were you with your performance in accomplishing these goals?

Good   Poor

16. Read the question description, and then write an "X" on the gradation line that corresponds to your answer.

FRUSTRATION LEVEL How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

FRUSTRATION

Low   High

17. Read the question description, and then write an "X" on the gradation line that corresponds to your answer.

TEMPORAL DEMAND How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was that pace slow and leisurely or rapid and frantic?

TEMPORAL DEMAND

Low   High

18. Answer only the question (A or B) that corresponds to your experiment.

(A) If your articles were presented <u>electronically</u>:

I would have preferred to read the articles on paper (circle one):   YES   NO

(B) If your articles were presented <u>on paper</u>:

I would have preferred to read the articles electronically (circle one): YES   NO

APPENDIX E: SHEET DESIGNATED FOR PROMPT RESPONSES AND NOTE-TAKING (NON-METACOGNITIVE)

Note-taking space:
Response to Prompt 1:
Response to Prompt 2:
Response to Prompt 3:

Passage from www.englishforeveryone.org

Please respond to all prompts and take any notes only on the sheet designated for prompt responses and note-taking.

Note: The comprehension test will consist of questions on factual recall as well as inferences.

Please start timer when ready to begin Prompt 1 below.

PROMPT 1: What do you know about the following topics (questionable treatment of animals, diseases, environmental impacts, and regulations) regarding animal factory farms? On the separate sheet designated for prompt responses and note-taking (in the Prompt 1 box), fill in – under the provided headings – any previous knowledge, experiences, or examples, you have related to these topics.

In addition (in the top-right corner of the Prompt 1 box), write a number 1-10 that indicates how difficult you believe it will be to learn about animal factory farms (1 = no problems at all, 10 = as difficult as material from the toughest class you have taken).

When finished with Prompt 1 and ready to start reading, please write the time showing on the timer:

Animal factory farms

According to some estimates, if we could compile the amount of food, land, water, and energy used to raise the 10 billion animals slaughtered each year for meat, we could use those resources to feed every single starving person on earth. The majority of these resources are depleted by concentrated animal feeding operations (CAFOs). CAFOs are factory farms that mass-produce livestock – harming animals, the environment, and humans in the process. It is true that these farming methods provide an abundant source of food and employ thousands of workers across the country. However, CAFOs should be placed under more stringent restrictions because of their unfair treatment of animals and the harm they do to both the environment and humans.

(paragraph 2) One of the key controversies surrounding factory farms is animal rights. Factory farms raise livestock indoors, as opposed to allowing the animals to graze in fields and pastures. The farmers favor this overcrowded environment because it maximizes profits. Providing less space for the animals costs less money; filling pens to their maximum capacity ensures that no space is wasted. Consequently, animal pens are often so small that larger animals cannot lie down or turn around. In some cases, these small cages are beneficial for more than just maximizing capacity: calves, for example, do not gain muscle mass in this environment. This keeps their meat more tender, which makes it more attractive to consumers.

(paragraph 3) Livestock in CAFOs are often found living in their own urine and feces, stimulating the spread of diseases—such as avian flu, foot and mouth disease, and mad cow disease—among other animals on the farm. In order to combat this, farmers must give the animals antibiotics. In many cases, however, antibiotics are used for disease prevention instead of treatment. In addition to being used to combat the spread of disease, antibiotics are also commonly used to encourage faster growth in

livestock. This overuse increases the risk of livestock developing immunity to antibiotics, ironically making animals even more susceptible to disease. After being digested, these antibiotics are released back into the environment in the form of milk, meat, and waste, which can affect the people who eat these products or the environment that absorbs them.

CAFOs also negatively impact the environment in the form of air and water pollution. Factory farms contribute to air pollution issues in the United States through the release of toxic gases and vapors and by burning fossil fuels to run farm machinery. These farms also have notable consequences for the environment in terms of water pollution. One characteristic of CAFOs that creates water pollution is the presence of a lagoon. Lagoons are artificial storage basins where animal excrement is temporarily contained; periodically, farmers flush this waste into ditches or nearby bodies of water. This waste combines with runoff from fertilized fields to pollute the water sources surrounding CAFOs. It adds excess nutrients, pathogens, veterinary pharmaceuticals, heavy metals, and excreted hormones to the water sources. Such pollutants not only affect aquatic life, but can lead to severe impacts on human health.

Another negative environmental impact of factory farms is resource depletion. Factory farming uses more land than any other agricultural or industrial enterprise in the country. CAFOs consume a great deal of resources in terms of grain, energy, and land. There are far more efficient ways of using these resources to feed people. For example, it has been estimated that the grain used to feed livestock in the United States alone could feed up to 800 million people in one year. By contrast, the production of livestock in CAFOs is a wasteful use of energy. While both chicken meat and soybeans are good sources of protein, producing equivalent amounts of protein from chicken meat and soybeans does not require equivalent amounts of energy: chicken meat production consumes 14 times more energy than soybean production. Grain and energy supplies should be used more efficiently to produce food sources other than livestock.

When finished with the text up until this point and ready to begin Prompt 2, please write the time showing on the timer:

PROMPT 2: On the separate sheet designated for prompt responses and note-taking (in the Prompt 2 box), check your understanding so far by creating flow maps of these two processes: How animal meat becomes tender (paragraph 2) and how antibiotics get into our food (paragraph 3). Refer back to text and concept maps review sheet as needed.

In addition (in the top-right corner of the Prompt 2 box), write a number 1-10 that reflects your understanding of the material so far (1 = have understood none of the material presented so far, 10 = all material so far is understood).

When finished with Prompt 2 and ready to continue reading, please write the time showing on the timer:

In order to combat the unfair treatment of animals and the risks to environmental and human health, CAFOs should be placed under stricter guidelines. One such regulation would force factory farms to adhere to air and water quality protection standards from which they have previously been exempted, like those set forth by the Clean Water Act. Enforcing these standards would lead to banning environmental hazards such as waste lagoons, which in turn would reduce environmental pollution and human health liabilities.

Some have suggested that due to these environmental and human health concerns, factory farms should be banned outright. Advocates for CAFOs, however, argue that factory farming allows for lower production costs that translate into lower food prices for consumers. Organic and free-range products, they argue, do not allow for the large-scale production of livestock; prices for meat, eggs, and dairy would increase should the country shift towards organic products. Although this would be an inconvenience to consumers, a price increase would encourage people to eat a diet of less meat. This cultural change would assist in solving the broader resource crisis as fewer grain, energy, and land resources would be needed to support smaller-scale production. Better treatment of animals and more responsible environmental practices would protect humans more from infectious diseases and the effects of air and water pollution—a benefit everyone should embrace.

When finished with the text up until this point and ready to begin Prompt 3, please write the time showing on the timer:

PROMPT 3: On the separate sheet designated for prompt responses and note-taking (in the Prompt 3 box), integrate your previous knowledge/experience with the information you learned from the article in a summary paragraph for each heading, and be sure to correct any misconceptions you might have had before reading the article (the box has the same headings as the one you filled in before reading the article).

In addition (in the top-right corner of the Prompt 3 box), write a number 1-10 that reflects how much your reading strategies (while reading this passage and answering prompts) deviated from what you usually do when reading (1 = no deviation at all, 10 = complete change from usual reading).

When finished with Prompt 3 and ready to consider the example questions, please write the time showing on the timer:

Use the example multiple-choice and short-response questions below to gauge your comprehension. These questions below do not have to be answered in any "official" manner; they are just for reference.

Multiple-choice: Which of the following pieces of information, if true, would best strengthen the author's argument in paragraph 6?

    a. People who reduce their intake of meat raised on factory farms reduce their chances of developing high blood pressure and heart conditions.
    b. A town in North Carolina that used to be situated next to a hog waste lagoon reported 50% fewer cases of respiratory illness after the lagoon was shut down.
    c. Factory farms that are forced to adhere to air and water quality protection standards often report a 30% decrease in annual profits.
    d. Several animal rights groups have supported bills to force factory farms to abide by greater environmental protection standards.
    e. The Clean Water Act was enacted in 1948 and expanded in 1972.

Short response: Does the author appear to be someone who would be in favor of "free-range" farming, in which livestock is raised outside? Defend your answer with two pieces of information from the article.

When you feel like you are ready to take the comprehension test, please STOP YOUR TIMER and write the time showing on the timer:

Alert the research assistant to proceed to the comprehension test.

Passage from www.englishforeveryone.org

Please respond to all prompts and take any notes only on the sheet designated for prompt responses and note-taking.

Note: The comprehension test will consist of questions on factual recall as well as inferences.

Please start timer when ready to begin Prompt 1 below.

Do you believe animals have emotions and feel pain like humans do? On the separate sheet designated for prompt responses and note-taking (Prompt 1), write 4-5 sentences defending your answer.

When finished with Prompt 1 and ready to start reading, please write the time showing on the timer:

Animal factory farms

According to some estimates, if we could compile the amount of food, land, water, and energy used to raise the 10 billion animals slaughtered each year for meat, we could use those resources to feed every single starving person on earth. The majority of these resources are depleted by concentrated animal feeding operations (CAFOs). CAFOs are factory farms that mass-produce livestock – harming animals, the environment, and humans in the process. It is true that these farming methods provide an abundant source of food and employ thousands of workers across the country. However, CAFOs should be placed under more stringent restrictions because of their unfair treatment of animals and the harm they do to both the environment and humans.

One of the key controversies surrounding factory farms is animal rights. Factory farms raise livestock indoors, as opposed to allowing the animals to graze in fields and pastures. The farmers favor this overcrowded environment because it maximizes profits. Providing less space for the animals costs less money; filling pens to their maximum capacity ensures that no space is wasted. Consequently, animal pens are often so small that larger animals cannot lie down or turn around. In some cases, these small cages are beneficial for more than just maximizing capacity: calves, for example, do not gain muscle mass in this environment. This keeps their meat more tender, which makes it more attractive to consumers.

Livestock in CAFOs are often found living in their own urine and feces, stimulating the spread of diseases—such as avian flu, foot and mouth disease, and mad cow disease—among other animals on the farm. In order to combat this, farmers must give the animals antibiotics. In many cases, however, antibiotics are used for disease prevention instead of treatment. In addition to being used to combat the spread of disease, antibiotics are also commonly used to encourage faster growth in livestock. This overuse increases the risk of livestock developing immunity to antibiotics, ironically making animals even more susceptible to disease. After being digested, these antibiotics are released back into the environment in the form of milk, meat, and waste, which can affect the people who eat these products or the environment that absorbs them.

CAFOs also negatively impact the environment in the form of air and water pollution. Factory farms contribute to air pollution issues in the United States through the release of toxic gases and vapors and

by burning fossil fuels to run farm machinery. These farms also have notable consequences for the environment in terms of water pollution. One characteristic of CAFOs that creates water pollution is the presence of a lagoon. Lagoons are artificial storage basins where animal excrement is temporarily contained; periodically, farmers flush this waste into ditches or nearby bodies of water. This waste combines with runoff from fertilized fields to pollute the water sources surrounding CAFOs. It adds excess nutrients, pathogens, veterinary pharmaceuticals, heavy metals, and excreted hormones to the water sources. Such pollutants not only affect aquatic life, but can lead to severe impacts on human health.

Another negative environmental impact of factory farms is resource depletion. Factory farming uses more land than any other agricultural or industrial enterprise in the country. CAFOs consume a great deal of resources in terms of grain, energy, and land. There are far more efficient ways of using these resources to feed people. For example, it has been estimated that the grain used to feed livestock in the United States alone could feed up to 800 million people in one year. By contrast, the production of livestock in CAFOs is a wasteful use of energy. While both chicken meat and soybeans are good sources of protein, producing equivalent amounts of protein from chicken meat and soybeans does not require equivalent amounts of energy: chicken meat production consumes 14 times more energy than soybean production. Grain and energy supplies should be used more efficiently to produce food sources other than livestock.

When finished with the text up until this point and ready to begin Prompt 2, please write the time showing on the timer:

Should the United States government be doing more to combat environmental pollution, and do you think pollution contributes to global warming? On the separate sheet designated for prompt responses and note-taking (Prompt 2), write 4-5 sentences defending your answer.

When finished with Prompt 2 and ready to continue reading, please write the time showing on the timer:

In order to combat the unfair treatment of animals and the risks to environmental and human health, CAFOs should be placed under stricter guidelines. One such regulation would force factory farms to adhere to air and water quality protection standards from which they have previously been exempted, like those set forth by the Clean Water Act. Enforcing these standards would lead to banning environmental hazards such as waste lagoons, which in turn would reduce environmental pollution and human health liabilities.

Some have suggested that due to these environmental and human health concerns, factory farms should be banned outright. Advocates for CAFOs, however, argue that factory farming allows for lower production costs that translate into lower food prices for consumers. Organic and free-range products, they argue, do not allow for the large-scale production of livestock; prices for meat, eggs, and dairy would increase should the country shift towards organic products. Although this would be an inconvenience to consumers, a price increase would encourage people to eat a diet of less meat. This cultural change would assist in solving the broader resource crisis as fewer grain, energy, and land resources would be needed to support smaller-scale production. Better treatment of animals and more responsible environmental practices would protect humans more from infectious diseases and the effects of air and water pollution—a benefit everyone should embrace.

When finished with the text up until this point and ready to begin Prompt 3, please write the time showing on the timer:

Do you believe that planet Earth is in danger of being depleted of natural resources in the future, and how important do you think it is for humans to find alternative sources of energy? On the separate sheet designated for prompt responses and note-taking (Prompt 3), write 4-5 sentences defending your answer.

When finished with Prompt 3 and ready to consider the example questions, please write the time showing on the timer:

Use the example multiple-choice and short-response questions below to gauge your comprehension. These questions below do not have to be answered in any "official" manner; they are just for reference.

Multiple-choice: Which of the following pieces of information, if true, would best strengthen the author's argument in paragraph 6?

a. People who reduce their intake of meat raised on factory farms reduce their chances of developing high blood pressure and heart conditions.
b. A town in North Carolina that used to be situated next to a hog waste lagoon reported 50% fewer cases of respiratory illness after the lagoon was shut down.
c. Factory farms that are forced to adhere to air and water quality protection standards often report a 30% decrease in annual profits.
d. Several animal rights groups have supported bills to force factory farms to abide by greater environmental protection standards.
e. The Clean Water Act was enacted in 1948 and expanded in 1972.

Short response: Does the author appear to be someone who would be in favor of "free-range" farming, in which farmers raise livestock outside? Defend your answer with two pieces of information from the article.

When you feel like you are ready to take the comprehension test, please STOP YOUR TIMER and write the time showing on the timer:

Alert the research assistant to proceed to the comprehension test.

APPENDIX H: PASSAGE 1 COMPREHENSION TEST – ANIMAL FACTORY FARMS

Animal factory farms – Comprehension test

Multiple-choice questions:

1. The primary purpose of the passage is to:

   a. persuade readers that factory farms should be more strictly regulated to minimize the harm they cause
   b. suggest economic alternatives to factory farms, such as organic farming and soybean production
   c. complain about the water pollution caused by the irresponsible practice of keeping waste lagoons on CAFOs
   d. educate readers about the pros and cons of CAFOs
   e. argue against the use of antibiotics for disease prevention in animals

2. Based on information in the passage, it can be inferred that animals raised on CAFOs live indoors because

   I.   animals that live indoors require fewer antibiotics than animals raised outside
   II.   some animals raised inside produce more appealing meat
   III. animals raised indoors are less expensive to maintain

   a. I only
   b. II only
   c. I and II only
   d. II and III only
   e. I, II, and III

3. Which of the following statements contains a valid objection to the author's argument in the final paragraph that he or she does not address?

   a. Not everyone wants to be a vegetarian.
   b. An increase in food prices is more than just an inconvenience for many households; it can mean the difference between having enough to eat or not.
   c. Factory farm owners pay taxes just like everyone else, and therefore they should be allowed input in the creation and modification of environmental policy.
   d. If farms move toward producing more organic products, food prices will likely go up.
   e. It is illegal to ban factory farms on the basis of current laws.

4. In the final paragraph, the author states that meat from animals raised on factory farms is cheaper than organic or free-range meat. The author likely admits this fact in order to

    a. provide a balanced perspective before concluding that the drawbacks of CAFO-produced meat outweigh the economic benefits
    b. give readers enough information so that they can come to their own conclusions about CAFOs' benefits and drawbacks
    c. warn readers about the economic dangers that would result from shutting down CAFOs
    d. criticize supporters of CAFOs for being more concerned with the economy than the environment
    e. argue that it is more important to address the broader resource crisis than to worry about food prices

5. The author describes CAFOs as having a negative impact on the environment for all of the following reasons except

    a. animal waste pollutes water sources near factory farms
    b. exhaust from farm machinery contributes to air pollution
    c. animals that live in overcrowded environments waste energy
    d. fertilizer-rich runoff from farms contaminates the environment
    e. antibiotics enter the human food supply through meat and milk

----

Short-response questions:

How would you describe the general tone of the article? Remember to defend your answer with specific examples of arguments from the author.

List and explain all of the possible effects of a move to free-range and organic products (positive or negative) and away from CAFOs. Write as much as you can remember.

Fast food plays a sizable role in driving American healthcare costs to unsustainable levels and its workers are often paid an unlivable minimum wage. Based on information from this article, do you think the author would be in favor of raising the minimum wage, which would necessarily cause an increase in fast food prices? Base your answer solely in the context of healthcare and fast food.

If you are done with these essay questions, raise your hand to let the research assistant know that you are ready to continue to the next article (if you have completed all three articles, you will move onto the survey).

Note-taking space:

Response to Prompt 1:   Difficulty rating (1-10): _____

    I.   Questionable treatment of animals on factory farms

    II.   Diseases on factory farms

    III.   Environmental impacts of factory farms

    IV.   Factory farm regulations

Response to Prompt 2:

Understanding (1-10): _____

How animal meat becomes tender

```
┌──────────┐     ┌──────────┐     ┌──────────┐     ┌──────────┐
│          │ ──> │          │ ──> │          │ ──> │          │
│          │     │          │     │          │     │          │
└──────────┘     └──────────┘     └──────────┘     └──────────┘
```

How antibiotics get into our food

```
┌──────────┐     ┌──────────┐     ┌──────────┐     ┌──────────┐
│          │ ──> │          │ ──> │          │ ──> │          │
│          │     │          │     │          │     │          │
└──────────┘     └──────────┘     └──────────┘     └──────────┘
```

Response to Prompt 3:

Deviation from usual strategies (1-10): _____

I.   Questionable treatment of animals on factory farms

II.   Diseases on factory farms

III.   Environmental impacts of factory farms

IV.   Factory farm regulations

Please respond to all prompts and take any notes only on the sheet designated for prompt responses and note-taking.

Note: The comprehension test will consist of questions on factual recall as well as inferences.

Please start timer when ready to begin Prompt 1 below.

What do you know about the following topics (symptoms and effects of PTSD, reasons why women are more likely to develop PTSD and have more severe symptoms than men, and obstacles to quality treatment for women with PTSD) regarding post-traumatic stress disorder? On the separate sheet designated for prompt responses and note-taking (in the Prompt 1 box), fill in – under the provided headings – any previous knowledge, experiences, or examples, you have related to these topics.

In addition (in the top-right corner of the Prompt 1 box), write a number 1-10 that indicates how difficult you believe it will be to learn about PTSD (1 = no problems at all, 10 = as difficult as material from the toughest class you have taken).

When finished with Prompt 1 and ready to start reading, please write the time showing on the timer:

Women and PTSD

Post-traumatic stress disorder (PTSD) is a severe anxiety disorder that affects millions of people around the world. Individuals can develop PTSD after experiencing any event that results in psychological trauma. Symptoms of PTSD involve flashbacks to the traumatic event, nightmares, obsessive behavior, anger, insomnia, difficulty concentrating, and hypervigilance. Individuals who suffer from PTSD can experience significant difficulties in social relationships, have lower self-esteem, and have trouble maintaining employment. People with PTSD experience a higher risk of committing suicide, developing a drug addiction, and suffering from alcoholism. Although PTSD can affect any individual, military veterans are especially susceptible to this debilitating affliction. Furthermore, within this population, women are more than twice as likely as men to develop PTSD. Studies have also shown that former service women who do develop PTSD experience more severe symptoms than their male counterparts. Recent changes made by the United States Department of Veterans Affairs have improved treatment options for female veterans living with PTSD, but there is still more that needs to be done.

Although studies investigating precisely why women are more likely than men to experience PTSD have not yet been completed, some experts have theorized that low unit cohesion is a major factor. Unit cohesion, defined as the mutual bonds of friendship and support among members of a military unit, is thought to be helpful in reducing the incidence of developing PTSD. According to recent surveys compiled by U.S. Army researchers, increased unit cohesion emerged as the most important factor determining whether soldiers developed suicidal thoughts. Women are more likely than men to experience low unit cohesion for a variety of reasons. One of the most obvious factors is the relative paucity of females in the military; currently, women make up only 20% of the armed forces. Pervasive male prejudice against women is another factor that can diminish unit cohesion for female soldiers. Because women are less likely than men to experience unit cohesion while serving in the military,

women are less likely to develop the social support structures that will help prevent them from developing PTSD, depression, or other serious mental health problems.

Another issue at play is the stigma amongst military personnel that asking for help for mental health issues makes one "weak." A recent Department of Defense study of returning combat troops shows that only 1 in 6 veterans acknowledged themselves to be suffering from symptoms of PTSD, and 3 out of 5 veterans were convinced that their comrades and commanding officers would lose confidence in them if they sought treatment for mental health issues. For women, this hesitation to self-identify as a sufferer of PTSD could be even greater; historically, female soldiers have struggled to be counted as equals to men on the battlefield. Women, stereotypically considered to possess less emotional fortitude than men, may be unwilling to admit that they are suffering from PTSD lest they appear to conform to this stereotype. Unfortunately for those who do not seek help, when PTSD goes untreated it is very likely to worsen over time.

When finished with the text up until this point and ready to begin Prompt 2, please write the time showing on the timer:

On the separate sheet designated for prompt responses and note-taking (in the Prompt 2 box), check your understanding of how women are different from men in terms of PTSD likelihood, PTSD severity, unit cohesion, and asking for help, by constructing bar graphs that qualitatively (roughly, not exactly) show the differences between the two groups. Refer back to text and concept maps review sheet as needed.

In addition (in the top-right corner of the Prompt 2 box), write a number 1-10 that reflects your understanding of the material so far (1 = have understood none of the material presented so far, 10 = all material so far is understood).

When finished with Prompt 2 and ready to continue reading, please write the time showing on the timer:

Another challenge is that until very recently, treatment for PTSD has been more difficult for women than men to obtain. Before rule changes were enacted in 2010, only veterans who encountered direct combat experience qualified to receive disability payments for PTSD. Because very few women are placed on the front lines, very few were eligible to receive free treatment for PTSD. However, recent regulation changes have ended these stipulations, allowing women who serve in any capacity to be eligible for benefits.

Even if female veterans are eligible for these benefits, the quality of the care a wartime PTSD sufferer receives can vary widely. The United States Department of Veterans Affairs (VA) pays disability benefits to service men and women who have been diagnosed with PTSD and also provides these individuals with free health care. But while mental health counseling that comes directly from VA doctors is completely free to veterans, there are often long waiting lists for those who need to be evaluated or treated. Therapy provided by non-VA professionals may not be covered by health insurance. Access to mental health professionals who have been specially trained to treat wartime PTSD is often difficult for those not living near major urban centers.

Perhaps the biggest impediment to achieving quality treatment for women suffering from wartime PTSD is a lack of research. While both the VA and independent agencies have completed hundreds of studies

researching the prevention of and treatment for PTSD amongst general military populations, there have been no studies completed that solely target women. Before adequate care can be provided, there must be greater understanding about the root causes of this issue as it affects women specifically. On both the research level and the policy level, more must be done to help the women who have sacrificed so much for their country.

When finished with the text up until this point and ready to begin Prompt 3, please write the time showing on the timer:

On the separate sheet designated for prompt responses and note-taking (in the Prompt 3 box), <u>integrate</u> your previous knowledge/experience with the information you learned from the article in a summary paragraph for each heading, and be sure to correct any misconceptions you might have had before reading the article (the box has the same headings as the one you filled in before reading the article).

In addition (in the top-right corner of the Prompt 3 box), write a number 1-10 that reflects how much your reading strategies (while reading this passage and answering prompts) deviated from what you usually do when reading (1 = no deviation at all, 10 = complete change from usual reading).

When finished with Prompt 3 and ready to consider the example questions, please write the time showing on the timer:

Use the example multiple-choice and short-response questions below to gauge your comprehension. These questions below do not have to be answered in any "official" manner; they are just for reference.

Multiple-choice: Based on information in the passage, it can be inferred that the author is most likely someone who

      a. does not believe women should be allowed to serve in the military
      b. is a mental health care professional
      c. has conducted studies on PTSD among veterans
      d. believes that advocating for others can lead to change
      e. is female

Short response: Explain whether an increase in the number of women in the military would, according to the author, help lower PTSD incidences among women.

When you feel like you are ready to take the comprehension test, please STOP YOUR TIMER and write the time showing on the timer:

Alert the research assistant to proceed to the comprehension test.

APPENDIX K: PASSAGE 2 – WOMEN AND PTSD (NON-METACOGNITIVE CONDITIONS)

Please respond to all prompts and take any notes only on the sheet designated for prompt responses and note-taking.

Note: The comprehension test will consist of questions on factual recall as well as inferences.

Please start timer when ready to begin Prompt 1 below.

Have you ever had a traumatic experience in your life that still comes back to bother you from time to time? On the separate sheet designated for prompt responses and note-taking (Prompt 1), write 4-5 sentences about that experience and how it changed you.

When finished with Prompt 1 and ready to start reading, please write the time showing on the timer:

Women and PTSD

Post-traumatic stress disorder (PTSD) is a severe anxiety disorder that affects millions of people around the world. Individuals can develop PTSD after experiencing any event that results in psychological trauma. Symptoms of PTSD involve flashbacks to the traumatic event, nightmares, obsessive behavior, anger, insomnia, difficulty concentrating, and hypervigilance. Individuals who suffer from PTSD can experience significant difficulties in social relationships, have lower self-esteem, and have trouble maintaining employment. People with PTSD experience a higher risk of committing suicide, developing a drug addiction, and suffering from alcoholism. Although PTSD can affect any individual, military veterans are especially susceptible to this debilitating affliction. Furthermore, within this population, women are more than twice as likely as men to develop PTSD. Studies have also shown that former service women who do develop PTSD experience more severe symptoms than their male counterparts. Recent changes made by the United States Department of Veterans Affairs have improved treatment options for female veterans living with PTSD, but there is still more that needs to be done.

Although studies investigating precisely why women are more likely than men to experience PTSD have not yet been completed, some experts have theorized that low unit cohesion is a major factor. Unit cohesion, defined as the mutual bonds of friendship and support among members of a military unit, is thought to be helpful in reducing the incidence of developing PTSD. According to recent surveys compiled by U.S. Army researchers, increased unit cohesion emerged as the most important factor determining whether soldiers developed suicidal thoughts. Women are more likely than men to experience low unit cohesion for a variety of reasons. One of the most obvious factors is the relative paucity of females in the military; currently, women make up only 20% of the armed forces. Pervasive male prejudice against women is another factor that can diminish unit cohesion for female soldiers. Because women are less likely than men to experience unit cohesion while serving in the military, women are less likely to develop the social support structures that will help prevent them from developing PTSD, depression, or other serious mental health problems.

Another issue at play is the stigma amongst military personnel that asking for help for mental health issues makes one "weak." A recent Department of Defense study of returning combat troops shows that only 1 in 6 veterans acknowledged themselves to be suffering from symptoms of PTSD, and 3 out of 5 veterans were convinced that their comrades and commanding officers would lose confidence in them if

they sought treatment for mental health issues. For women, this hesitation to self-identify as a sufferer of PTSD could be even greater; historically, female soldiers have struggled to be counted as equals to men on the battlefield. Women, stereotypically considered to possess less emotional fortitude than men, may be unwilling to admit that they are suffering from PTSD lest they appear to conform to this stereotype. Unfortunately for those who do not seek help, when PTSD goes untreated it is very likely to worsen over time.

When finished with the text up until this point and ready to begin Prompt 2, please write the time showing on the timer:

What other jobs do you know of where women are stigmatized against? On the separate sheet designated for prompt responses and note-taking (Prompt 2), identify the job and explain the stigmas that women have to deal with on the job (4-5 sentences).

When finished with Prompt 2 and ready to continue reading, please write the time showing on the timer:

Another challenge is that until very recently, treatment for PTSD has been more difficult for women than men to obtain. Before rule changes were enacted in 2010, only veterans who encountered direct combat experience qualified to receive disability payments for PTSD. Because very few women are placed on the front lines, very few were eligible to receive free treatment for PTSD. However, recent regulation changes have ended these stipulations, allowing women who serve in any capacity to be eligible for benefits.

Even if female veterans are eligible for these benefits, the quality of the care a wartime PTSD sufferer receives can vary widely. The United States Department of Veterans Affairs (VA) pays disability benefits to service men and women who have been diagnosed with PTSD and also provides these individuals with free health care. But while mental health counseling that comes directly from VA doctors is completely free to veterans, there are often long waiting lists for those who need to be evaluated or treated. Therapy provided by non-VA professionals may not be covered by health insurance. Access to mental health professionals who have been specially trained to treat wartime PTSD is often difficult for those not living near major urban centers.

Perhaps the biggest impediment to achieving quality treatment for women suffering from wartime PTSD is a lack of research. While both the VA and independent agencies have completed hundreds of studies researching the prevention of and treatment for PTSD amongst general military populations, there have been no studies completed that solely target women. Before adequate care can be provided, there must be greater understanding about the root causes of this issue as it affects women specifically. On both the research level and the policy level, more must be done to help the women who have sacrificed so much for their country.

When finished with the text up until this point and ready to begin Prompt 3, please write the time showing on the timer:

What other issues or fields do you think need to be investigated more through research in this country? On the separate sheet designated for prompt responses and note-taking (Prompt 3), explain why more research needs to be done in that field or regarding that issue (4-5 sentences).

When finished with Prompt 3 and ready to consider the example questions, please write the time showing on the timer:

Use the example multiple-choice and short-response questions below to gauge your comprehension. These questions below do not have to be answered in any "official" manner; they are just for reference.

Multiple-choice: Based on information in the passage, it can be inferred that the author is most likely someone who

     a. does not believe women should be allowed to serve in the military
     b. is a mental health care professional
     c. has conducted studies on PTSD among veterans
     d. believes that advocating for others can lead to change
     e. is female

Short response: Explain whether an increase in the number of women in the military would, according to the author, help lower PTSD incidences among women.

When you feel like you are ready to take the comprehension test, please STOP YOUR TIMER and write the time showing on the timer:

Alert the research assistant to proceed to the comprehension test.

APPENDIX L: PASSAGE 2 COMPREHENSION TEST – WOMEN AND PTSD

Women and PTSD – Comprehension test

Multiple-choice questions:

1. The primary purpose of the passage is to:

    a. explain why military veterans are more likely than civilians to develop PTSD
    b. persuade government officials to increase funding for PTSD treatment centers in non-urban areas
    c. denounce the United States military for the way they have handled veterans' mental health problems
    d. inform readers about the likely warning signs of PTSD among military veterans
    e. educate readers about the problem of insufficient treatment available for female veterans with PTSD

2. According to the passage, all of the following are reasons why females develop PTSD more frequently than males except

    a. women experience lower unit cohesion than men
    b. women can face negative stereotypes in the military
    c. women are more likely than men to suffer from depression
    d. males in the military greatly outnumber females
    e. males tend to develop stronger social bonds during their service than females do

3. Based on information in the passage, it can be inferred that negative stereotypes about women in the military contribute to their increased likelihood to develop PTSD in which of the following ways?

    I.   Some male members of the armed forces subscribe to negative stereotypes about women; this prejudice may prevent women from forming close bonds with their units.
    II. Women may be less likely than men to admit to suffering from PTSD because they do not want to conform to stereotypes that portray women as weak.
    III. Women are aware of the negative stereotypes that pervade the military. This awareness may lead to a reduction in self-esteem.

    a. I only
    b. II only
    c. I and II only
    d. II and III only
    e. I, II, and III

4. Based on information in the passage, which of the following statements best reflects the author's opinion about the mental health care provided for female veterans?

    a. Although the VA has not done nearly enough, state and independent agencies have made up for government deficiencies.

    b. The government has done almost nothing to help; the way we treat our female veterans is a national disgrace.

    c. Because service women are more likely than service men to develop PTSD, mental health care has been better for female veterans.

    d. The VA and other government agencies have attempted to provide mental health care for female veterans, but most of them refuse to seek treatment.

    e. The amount of care provided has improved over the past few years, but it is still insufficient.

5. Which of the following pieces of evidence, if true, would best strengthen the author's assertion that PTSD care varies widely in quality?

    a. Since 2001, when operations in Iraq and Afghanistan began, more than 230,000 women have served in the United States Military.

    b. Studies have shown that those who receive therapy for PTSD are less likely to commit suicide and develop substance abuse problems.

    c. Women are 50% more likely to experience sexual assault while deployed than men.

    d. Although experts now argue that intensive cognitive therapy is more effective at treating wartime PTSD than prescription drugs, most mental health professionals are trained only to offer pharmaceutical solutions for PTSD.

    e. In addition to causing obvious mental and social health problems for those who suffer from the disorder, PTSD can also cost society up to $6.2 billion in the form of lost labor productivity and welfare expenses.

Short-response questions:

Identify another population group, besides the military, that you think might be especially susceptible to developing PTSD. Defend your answer with two facts from the article and discuss what parts of their jobs/lives might lead to PTSD.
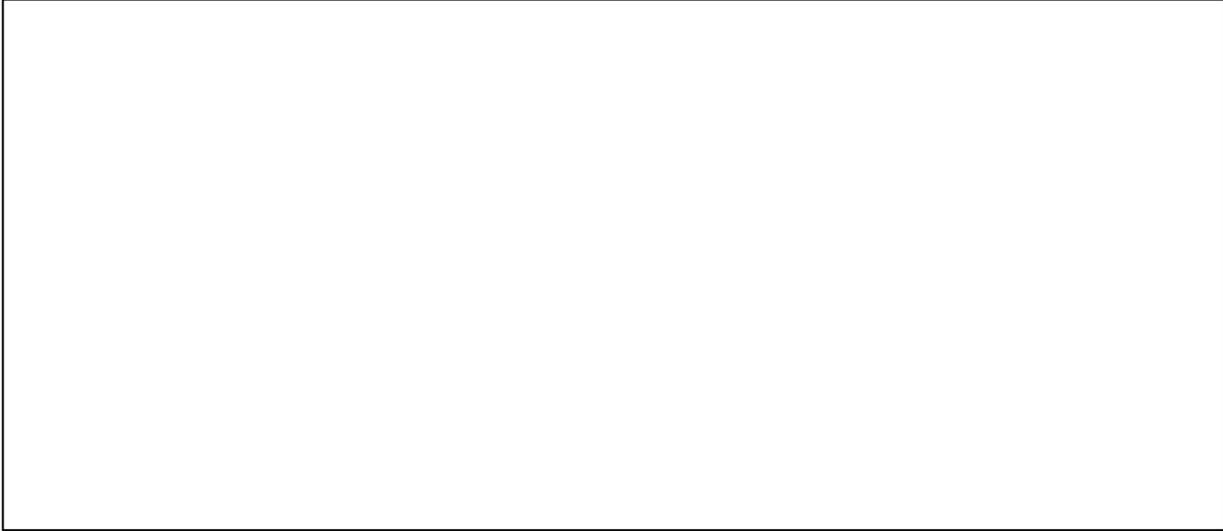
Explain what unit cohesion is and describe why women often experience lower levels of it than men in any given military unit.

The author is concerned about a lack of PTSD studies that target women. Describe a study (of existing data) or an experiment that could address any one of the issues outlined in the article. Include information about methodology and expected results.
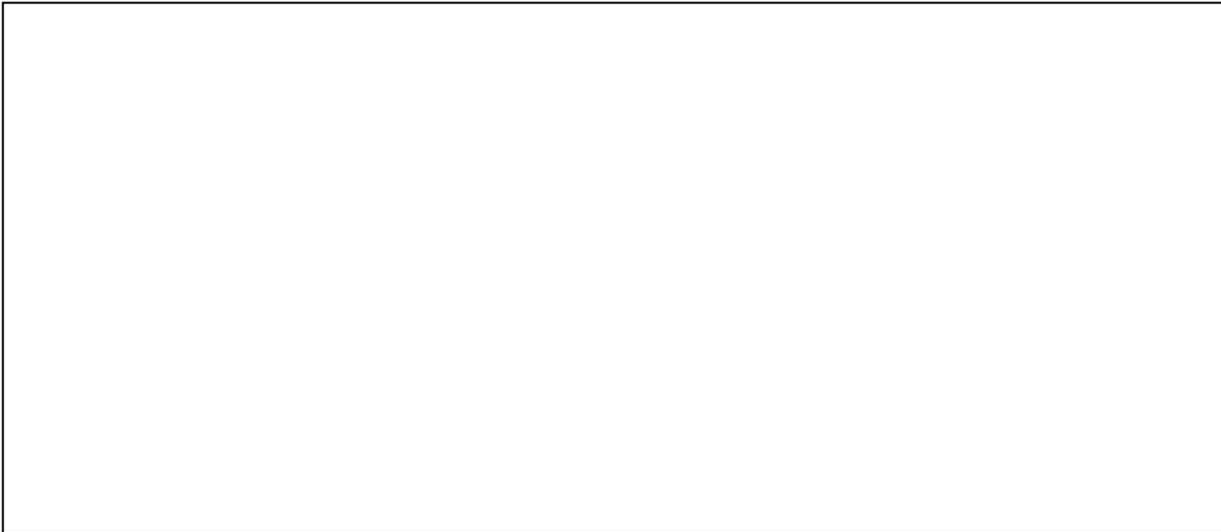
If you are done with these essay questions, raise your hand to let the research assistant know that you are ready to continue to the next article (if you have completed all three articles, you will move onto the survey).

APPENDIX M: PASSAGE 2 (WOMEN AND PTSD) SHEET DESIGNATED FOR PROMPT RESPONSES AND NOTE-TAKING (METACOGNITION CONDITIONS)

Note-taking space:

Response to Prompt 1:   Difficulty rating (1-10): _____

    V.   Symptoms and effects of PTSD

    VI.   Reasons why women are more likely to develop PTSD (and have more severe symptoms)

    VII.   Obstacles to quality treatment for women with PTSD

Response to Prompt 2:

Likelihood of developing PTSD

	Male	Female

Severity of PTSD symptoms

	Male	Female

Support from members of unit

	Male	Female
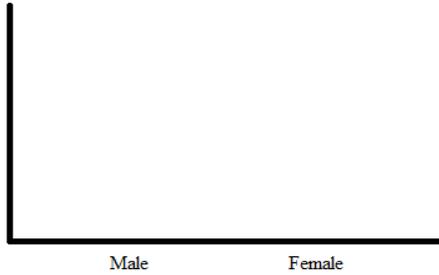
Likelihood of reporting PTSD symptoms

	Male	Female
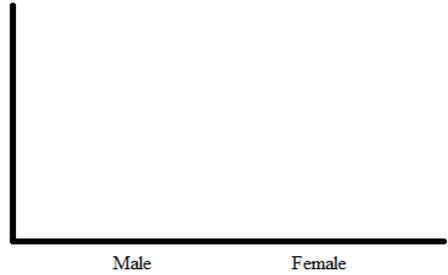
Response to Prompt 3:

Deviation from usual strategies (1-10): _____

I.   Symptoms and effects of PTSD

II.   Reasons why women are more likely to develop PTSD (and have more severe symptoms)

III.   Obstacles to quality treatment for women with PTSD

Please respond to all prompts and take any notes only on the sheet designated for prompt responses and note-taking.

Note: The comprehension test will consist of questions on factual recall as well as inferences.

Please start timer when ready to begin Prompt 1 below.

What do you know about the differences in culture, sound, undertone, and imagery, between words of very similar meaning (synonyms)? On the separate sheet designated for prompt responses and note-taking (in the Prompt 1 box), fill in – under the provided headings – any previous knowledge, experiences, or examples, you have related to these topics.

In addition (in the top-right corner of the Prompt 1 box), write a number 1-10 that indicates how difficult you believe it will be to learn about the above topics related to synonyms (1 = no problems at all, 10 = as difficult as material from the toughest class you have taken).

When finished with Prompt 1 and ready to start reading, please write the time showing on the timer:

Synonyms

It is not uncommon for close synonyms to be understood to share the same meaning. The difference between words like "hard" and "difficult", for example, goes tragically unnoticed. One may employ one or the other with complete indifference, postulating no discrepancy between them. In general this is well and good; most people lack the scrupulous pedanticalness to quibble over such trifles. Nevertheless, for those of us with ample compulsiveness (and time), it is of significant value to comprehend such nuances.

Take for example the following sentences: 1) The test was hard. 2) The test was difficult. Is the difference between these synonyms readily apparent? Is there a noticeable difference between them at all? Indeed, these questions are valid and warrant answer. For, what would be the point to having multiple words with the exact same meaning? No, that would be superfluous; the English language being far too economical. While many close synonyms share similar, if not the same, dictionary definitions, the feeling, or mood, they convey is utterly singular. Although a dictionary can provide information about word meanings, pronunciations, etymologies, inflected forms, derived forms, et cetera, it cannot communicate how it feels to use a word.

So, if there is indeed a difference between words like "hard" and "difficult", what is it? To begin, "hard" is pragmatic and realistic, firmly grounded in reality. It is a utilitarian word that gets the job done and doesn't apologize for its brusque, uncouth nature. On the other hand, "difficult" is eloquent and refined. It is civilized, willing to expend the effort necessary to appear urbane. Why, the mere difference in sonic quality between them is striking enough. "Hard" makes a quick, unassuming sound, having but a single syllable (voiced under certain inflections, it can even come across as harsh), while "difficult" is more lengthy and melodic, its number of syllables totaling three times that of its counterpart. Furthermore "hard" is more likely to be used in casual, informal circumstances, or to communicate an idea "on the go" or simply to "get it out" as the sayings go. It is used without pretense, and does not maintain a feeling of being overly concerned. In terms of daily usage, "hard" may be employed by an exhausted

brick mason when posed with the question, "How was your day?" Conversely, "difficult" may be used by a military general upon explaining to his or her superior the progression of a particularly taxing campaign.

Similar to "hard" and "difficult", the words "weird" and "strange" too are close synonyms, and may seemingly be used interchangeably. Take for instance the following sentences: 1) Sea monkeys are weird. 2) Sea monkeys are strange. Contrary to popular belief, these sentences are not tautologous. So how do they differ? Their dictionary definitions are nearly identical, so the difference does not lie there. Rather, the difference involves the feeling, or mood, that these words convey. Notice that while "weird" and "strange" both have but one syllable, the latter has a remarkably distinguished feel. Similar to "hard", "weird" conveys a more basic, a more crude, sentiment. Something "weird" is crass or gross, and is typically undesirable. No one wants to be associated with something "weird". If trying to impress someone, one probably doesn't want to be categorized among the "weird". On the other hand, if something is labeled as "strange", it is not necessarily bad. Rather, something "strange" is simply abnormal, or unusual—a deviation from what is expected. This distinction between "weird" and "strange" is so pronounced that the latter can be used as a euphemism for the former in certain situations. For example, notice how a simple substitution is able to make the following sentence less offensive: "Your mother's cookies taste weird" compared to "Your mother's cookies taste strange". In the former sentence, the speaker sounds as though he or she is insulting your mother's cookies, stating that they taste bad. In the latter sentence, however, the speaker sounds as though the cookies simply taste different, or unusual, compared to what he or she is used to—the difference owing to the innocuous addition of too much flour, perhaps.

When finished with the text up until this point and ready to begin Prompt 2, please write the time showing on the timer:

On the separate sheet designated for prompt responses and note-taking (in the Prompt 2 box), check your understanding so far by filling in the provided "double bubble" map to compare and contrast two synonyms on some of the criteria from the article or some that you can think of (use as many of the bubbles as you can). Possible word pairs (if you can't think of any): Beautiful and stunning, brilliant and smart, rich and wealthy, mean and unpleasant. Refer back to text and concept maps review sheet as needed.

In addition (in the top-right corner of the Prompt 2 box), write a number 1-10 that reflects your understanding of the material so far (1 = have understood none of the material presented so far, 10 = all material so far is understood).

When finished with Prompt 2 and ready to continue reading, please write the time showing on the timer:

Finally, let's look at the synonyms, "happy" and "glad". As in the aforementioned cases, these words seem to have little or no discernible difference between them. Take for example the following sentences: 1) Tommy is happy because he got a new bike. 2) Tommy is glad because he got a new bike. Most understand these sentences to have the same meaning. And again, upon consulting a dictionary, one will find highly similar, if not the same, definitions. But these definitions lack the feeling, the unique emotional charge that these words convey. The word "happy" conveys a sense of levity, or a carefree attitude. The thought of someone who is "happy" conjures the image of a bright-eyed, ruddy, smiling face. One is "happy" on the morning of his birthday, discovering a new puppy bounding into his

bedroom. On the other hand, the word "glad" conveys a sense of relief or contentment. The thought of someone who is "glad" conjures the image of a man standing crossed-armed, nodding gently, a stoic grin crossing his face. One is "glad" when he sees that the child's lost puppy has been found, and was merely frolicking too far from home.

Granted, the notion that close synonyms can be used interchangeably is prevalent among English speakers. And alas, the dictionary—the text purported to be responsible for clarifying such issues—is of little assistance. In the end, it is left to us, the speakers of the language, those actively responsible for maintaining its sustenance and generation, to understand how these words make us feel and what mood we are inclined to attach to them. Using the examples and insights described above, one may come to recognize these subtle, yet crucial, differences.

When finished with the text up until this point and ready to begin Prompt 3, please write the time showing on the timer:

On the separate sheet designated for prompt responses and note-taking (in the Prompt 3 box), <u>integrate</u> your previous knowledge/experience with the information you learned from the article in a summary paragraph for each heading, and be sure to correct any misconceptions you might have had before reading the article (the box has the same headings as the one you filled in before reading the article).

In addition (in the top-right corner of the Prompt 3 box), write a number 1-10 that reflects how much your reading strategies (while reading this passage and answering prompts) deviated from what you usually do when reading (1 = no deviation at all, 10 = complete change from usual reading).

When finished with Prompt 3 and ready to consider the example questions, please write the time showing on the timer:

<u>Use the example multiple-choice and short-response questions below to gauge your comprehension. These questions below do not have to be answered in any "official" manner; they are just for reference.</u>

Multiple-choice: Which of the following statements best describes the main idea of this passage?

> a. Close synonyms are difficult to comprehend, and are commonly used interchangeably.
> b. Contrary to popular belief, close synonyms do not share the same meaning.
> c. The difference between the words "hard" and "difficult" is indiscernible to most.
> d. Absent a dictionary definition, the difference between close synonyms is difficult to discern.
> e. Close synonyms can be ascribed their own individual feeling or mood.

Short response: What is the general relationship between a word's length and its perceived nature or feeling?

When you feel like you are ready to take the comprehension test, please STOP YOUR TIMER and write the time showing on the timer:

Alert the research assistant to proceed to the comprehension test.

Please respond to all prompts and take any notes only on the sheet designated for prompt responses and note-taking.

Note: The comprehension test will consist of questions on factual recall as well as inferences.

Please start timer when ready to begin Prompt 1 below.

What is your favorite book? On the separate sheet designated for prompt responses and note-taking (Prompt 1), write 4-5 sentences about what you liked about that book.

When finished with Prompt 1 and ready to start reading, please write the time showing on the timer:

Synonyms

It is not uncommon for close synonyms to be understood to share the same meaning. The difference between words like "hard" and "difficult", for example, goes tragically unnoticed. One may employ one or the other with complete indifference, postulating no discrepancy between them. In general this is well and good; most people lack the scrupulous pedanticalness to quibble over such trifles. Nevertheless, for those of us with ample compulsiveness (and time), it is of significant value to comprehend such nuances.

Take for example the following sentences: 1) The test was hard. 2) The test was difficult. Is the difference between these synonyms readily apparent? Is there a noticeable difference between them at all? Indeed, these questions are valid and warrant answer. For, what would be the point to having multiple words with the exact same meaning? No, that would be superfluous; the English language being far too economical. While many close synonyms share similar, if not the same, dictionary definitions, the feeling, or mood, they convey is utterly singular. Although a dictionary can provide information about word meanings, pronunciations, etymologies, inflected forms, derived forms, et cetera, it cannot communicate how it feels to use a word.

So, if there is indeed a difference between words like "hard" and "difficult", what is it? To begin, "hard" is pragmatic and realistic, firmly grounded in reality. It is a utilitarian word that gets the job done and doesn't apologize for its brusque, uncouth nature. On the other hand, "difficult" is eloquent and refined. It is civilized, willing to expend the effort necessary to appear urbane. Why, the mere difference in sonic quality between them is striking enough. "Hard" makes a quick, unassuming sound, having but a single syllable (voiced under certain inflections, it can even come across as harsh), while "difficult" is more lengthy and melodic, its number of syllables totaling three times that of its counterpart. Furthermore "hard" is more likely to be used in casual, informal circumstances, or to communicate an idea "on the go" or simply to "get it out" as the sayings go. It is used without pretense, and does not maintain a feeling of being overly concerned. In terms of daily usage, "hard" may be employed by an exhausted brick mason when posed with the question, "How was your day?" Conversely, "difficult" may be used by a military general upon explaining to his or her superior the progression of a particularly taxing campaign.

Similar to "hard" and "difficult", the words "weird" and "strange" too are close synonyms, and may seemingly be used interchangeably. Take for instance the following sentences: 1) Sea monkeys are

weird. 2) Sea monkeys are strange. Contrary to popular belief, these sentences are not tautologous. So how do they differ? Their dictionary definitions are nearly identical, so the difference does not lie there. Rather, the difference involves the feeling, or mood, that these words convey. Notice that while "weird" and "strange" both have but one syllable, the latter has a remarkably distinguished feel. Similar to "hard", "weird" conveys a more basic, a more crude, sentiment. Something "weird" is crass or gross, and is typically undesirable. No one wants to be associated with something "weird". If trying to impress someone, one probably doesn't want to be categorized among the "weird". On the other hand, if something is labeled as "strange", it is not necessarily bad. Rather, something "strange" is simply abnormal, or unusual—a deviation from what is expected. This distinction between "weird" and "strange" is so pronounced that the latter can be used as a euphemism for the former in certain situations. For example, notice how a simple substitution is able to make the following sentence less offensive: "Your mother's cookies taste weird" compared to "Your mother's cookies taste strange". In the former sentence, the speaker sounds as though he or she is insulting your mother's cookies, stating that they taste bad. In the latter sentence, however, the speaker sounds as though the cookies simply taste different, or unusual, compared to what he or she is used to—the difference owing to the innocuous addition of too much flour, perhaps.

When finished with the text up until this point and ready to begin Prompt 2, please write the time showing on the timer:

Have you ever said something to someone and wish you had phrased it or explained it differently? On the separate sheet designated for prompt responses and note-taking (Prompt 2), explain the situation, what you said, and what you wish you had said (4-5 sentences).

When finished with Prompt 2 and ready to continue reading, please write the time showing on the timer:

Finally, let's look at the synonyms, "happy" and "glad". As in the aforementioned cases, these words seem to have little or no discernible difference between them. Take for example the following sentences: 1) Tommy is happy because he got a new bike. 2) Tommy is glad because he got a new bike. Most understand these sentences to have the same meaning. And again, upon consulting a dictionary, one will find highly similar, if not the same, definitions. But these definitions lack the feeling, the unique emotional charge that these words convey. The word "happy" conveys a sense of levity, or a carefree attitude. The thought of someone who is "happy" conjures the image of a bright-eyed, ruddy, smiling face. One is "happy" on the morning of his birthday, discovering a new puppy bounding into his bedroom. On the other hand, the word "glad" conveys a sense of relief or contentment. The thought of someone who is "glad" conjures the image of a man standing crossed-armed, nodding gently, a stoic grin crossing his face. One is "glad" when he sees that the child's lost puppy has been found, and was merely frolicking too far from home.

Granted, the notion that close synonyms can be used interchangeably is prevalent among English speakers. And alas, the dictionary—the text purported to be responsible for clarifying such issues—is of little assistance. In the end, it is left to us, the speakers of the language, those actively responsible for maintaining its sustenance and generation, to understand how these words make us feel and what mood we are inclined to attach to them. Using the examples and insights described above, one may come to recognize these subtle, yet crucial, differences.

When finished with the text up until this point and ready to begin Prompt 3, please write the time showing on the timer:

Do you think the rise of texting (and other short-form communication) is leading to a general decline in the ability of younger people to write properly and with nuance? On the separate sheet designated for prompt responses and note-taking (Prompt 3), state whether you think so and defend your answer (4-5 sentences).

When finished with Prompt 3 and ready to consider the example questions, please write the time showing on the timer:

Use the example multiple-choice and short-response questions below to gauge your comprehension. These questions below do not have to be answered in any "official" manner; they are just for reference.

Multiple-choice: Which of the following statements best describes the main idea of this passage?

     a. Close synonyms are difficult to comprehend, and are commonly used interchangeably.
     b. Contrary to popular belief, close synonyms do not share the same meaning.
     c. The difference between the words "hard" and "difficult" is indiscernible to most.
     d. Absent a dictionary definition, the difference between close synonyms is difficult to discern.
     e. Close synonyms can be ascribed their own individual feeling or mood.

Short response: What is the general relationship between a word's length and its perceived nature or feeling?

When you feel like you are ready to take the comprehension test, please STOP YOUR TIMER and write the time showing on the timer:

Alert the research assistant to proceed to the comprehension test.

APPENDIX P: PASSAGE 3 COMPREHENSION TEST – SYNONYMS

Synonyms – Comprehension test

Multiple-choice questions:

1. What is the thesis statement in this passage?

   a. It is not uncommon for close synonyms to be believed to share the same meaning.
   b. However, for those of us with ample compulsiveness (and time), it is of significant value to comprehend such nuances.
   c. While many close synonyms share similar, if not the same, dictionary definitions, the feeling, or mood, they convey is utterly singular.
   d. Although a dictionary can provide information about word meanings, pronunciations, etymologies, inflected forms, derived forms, et cetera, it cannot communicate the energy of a word.
   e. In the end, it is left to us, the speakers of the language, those actively responsible for maintaining its sustenance and generation, to understand how these words make us feel and what mood we are inclined to attach to them.

2. Using the information from the passage, it can be inferred that which of the following statements contains a tautology?

   a. Paige received a free gift for her birthday.
   b. Science tells us that humans evolved to use their thumbs.
   c. Computers represent a significant technological advancement.
   d. Hexagons have six sides.
   e. Drugs are bad.

3. It can be inferred that the author believes there to exist a relationship between the vulgarity of a word and the

   I. number of syllables it has
   II. way it sounds
   III. way it is commonly used

   a. I only
   b. II only
   c. I and II only
   d. II and III only
   e. I, II, and III

4. With respect to the way in which close synonyms are commonly understood, the author's tone can best be described as

    a. belligerent
    b. supercilious
    c. rueful
    d. conscientious
    e. blasé

5. Which of the following, according to the author, does <u>not</u> change the "feeling" of a word?

    a. Perceived eloquence
    b. Speed with which word can be spoken
    c. Number of letters in the word
    d. Image conjured by the word
    e. Number of syllables in the word

Short-response questions:

If the author was in a car accident and narrowly avoided serious injury, would he use the word "happy" or "glad" to describe his emotions? Pick one of the words and defend your answer using information from the article.
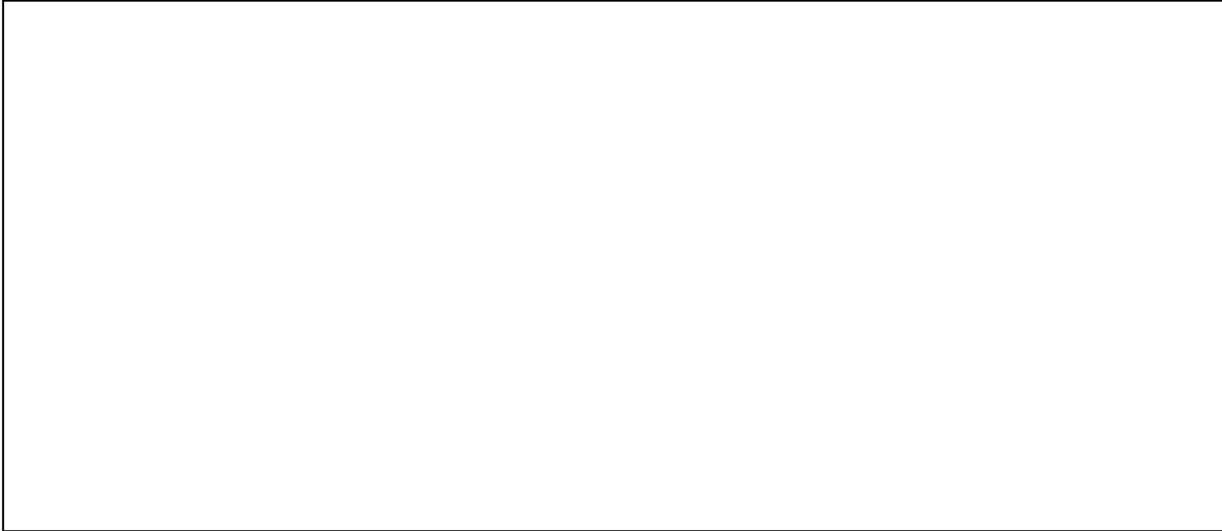
The words "nosy," "inquisitive," and "curious," have similar meanings but different connotations. Rank these words in order of how favorable they are in terms of describing a person and defend your ranking with arguments that the author uses in the passage.

Who is in the intended audience for this passage and what purpose(s) does the article serve for those people? Defend your answer with information from the article.

If you are done with these essay questions, raise your hand to let the research assistant know that you are ready to continue to the next article (if you have completed all three articles, you will move onto the survey).

Note-taking space:

Response to Prompt 1: | Difficulty rating (1-10): _____

VIII.   Differences between words in terms of culture

IX.   Differences between words in terms of sound

X.   Differences between words in terms of undertone

XI.   Differences between words in terms of imagery

Response to Prompt 2:
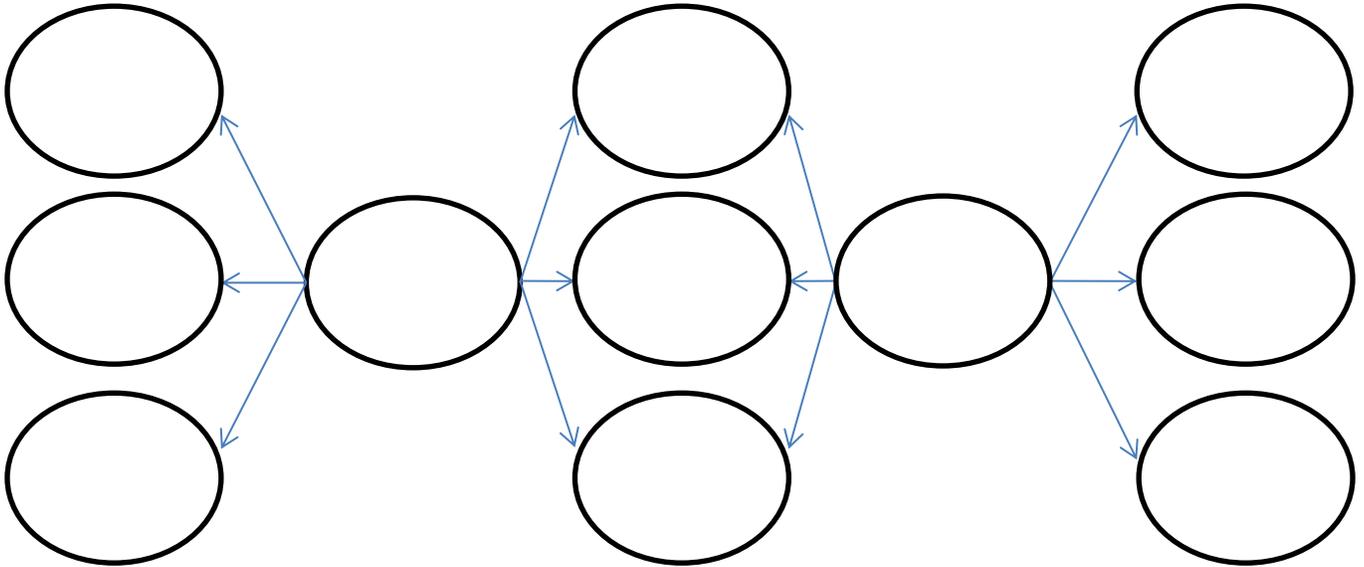
(nine empty ovals connected by blue arrows)

Response to Prompt 3:

    I.   Differences between words in terms of culture

    II.   Differences between words in terms of sound

    III.   Differences between words in terms of undertone

    IV.   Differences between words in terms of imagery

REFERENCES

Ackerman, R. & Goldsmith, M. (2011). Metacognitive regulation of text learning: On screen versus on paper. *Journal of Experimental Psychology: Applied*, *17* (1), 18-32.

Ackerman, R. & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior*, *28* (5), 1816-1828.

Aleven V. & Koedinger, K. R. (2000). Limitations of student control: Do students know when they need help? In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), Proceedings of the 5th International Conference on Intelligent Tutoring Systems, pp. 292-303. Amsterdam: IOS Press.

Andre, A. D. & Wickens, C. D. (1995, October). When users want what's not best for them. *Ergonomics in Design*, pp. 10-14.

Anuradha, K. T. & Usha, H. S. (2006). Use of e-books in an academic and research environment: A case study from the Indian Institute of Science. *Program: Electronic library and information systems*, *40* (1), 48-62.

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research*, *30* (1-2), 87-111.

Barkley, E. (2010). *Student engagement techniques: A handbook for college faculty*. San Francisco, CA: Jossey-Bass.

Beal, C. R., Qu, L., & Lee, H. (2006). Classifying learner engagement through integration of multiple data sources. *Proceedings of the 21st National Conference on Artificial Intelligence*, *1*, 151-156.

Bodmann, S. M. & Robinson, D. H. (2004). Speed and performance differences among computer-based and paper-pencil tests. *Journal of Educational Computing Research*, *31* (1), 51-60.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In: F. Reiner, & R. Kluwe (Eds.), *Metacognition, motivation, and understanding*. Hillsdale, NJ: Erlbaum.

Burleson, W. (2005). Developing creativity, motivation, and self-actualization with learning systems. *International Journal of Human-Computer Studies*, *63* (4-5), 436-451.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65,* 245-281.

Buzzetto-More, N., Sweat-Guy, R., Elobaid, M. (2007). Reading in a digital age: E-books. Are students ready for this learning object? *Interdisciplinary Journal of Knowledge and Learning Objects*, *3*, 239-250

Byrnes, J. P. (2003). Factors predictive of mathematics achievement in white, black, and Hispanic 12th graders. *Journal of Educational Psychology*, *95*, 316-326.

Creed, A., Dennis, I. & Newstead, S. (1987). Proof-reading on VDUs. *Behaviour and Information Technology*, *6* (1), 3-13.

Czaja, S. J. & Sharit, J. (1998). Age Differences in Attitudes Toward Computers. *Journal of Gerontology: Psychological Sciences, 53B* (5), 329-340.

DeAngelis, S. (2000). Equivalency of computer-based and paper-and-pencil testing. *Journal of Allied Health*, *29*, 161-164.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the Conceptual Lens on

Metacognition, Self-regulation, and Self-regulated Learning. *Educational Psychology*

Review, *20* (4), 391-409.

Dunlosky, J. & Thiede, K. W. (1998). What makes people study more? An evaluation of factors

that affect self-paced study. *Acta Psychologica, 98,* 37-56.

Dyson, M. & Haselgrove, M. (2000). The effects of reading speed and reading patterns on the

understanding of text read from screen. *Journal of Research in Reading*. *23* (2), 210-223.

Ertmer, P. A. & Newby, T. J. (1996). The expert learner: strategic, self-regulated, and reflective.

Instructional Science, *24*, 1–24.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L.B. Resnick (Ed.), *The*

nature of intelligence (pp. 231-235). Hillsdale, NJ: Erlbaum.

Gerlach, J. & Buxmann, P. (2011). Investigating the acceptance of electronic books − The impact

of haptic dissonance on innovation adoption. *Proceedings of the European Conference on*

Information Systems (Paper 141).

Gould, J. D., Alfaro, L., Barnes, V., Finn, R., Grischkowsky, N., & Minuto, A. (1987a). Reading

is slower from CRT displays than from paper: attempts to isolate a single-variable

explanation. *Human Factors*, *29* (3), 269-299.

Gould, J. D., Alfaro, L., Finn, R., Haupt, B., & Minuto, A. (1987b). Reading from CRT displays

can be as fast as reading from paper. *Human Factors*, *29* (5), 497-517.

Gould, J. D. & Grischkowsky, N. (1984). Doing the same work with hard copy and cathode-ray

tube (CRT) computer terminals. *Human Factors*, *26* (3), 323-337.

Horton, S.V. & Lovitt, T.C. (1994). A comparison of two methods of administrating group

reading inventories to diverse learners. *Remedial and Special Education*, *15*, 378–390.

Jabr, F. (2013, April 11). *The Reading Brain In the Digital Age: The Science of Paper Versus Screens*. Retrieved from http://www.scientificamerican.com/article/reading-paper-screens/

Jamali, H. R., Nicholas, D., & Rowlands, I. (2009). Scholarly e-books: The views of 16,000 academics: Results from the JLSC National E-Book Observatory. *Aslib Proceedings: New Information Perspectives*, *61*, 33047.

Kauffmann, D. F., Ge, X., Xie, K., & Chen, C. H. (2008). Prompting in web-based environments: Supporting self-monitoring and problem solving skills in college students. *Journal of Educational Computing Research*, *38* (2), 115-137

Kauffmann, D. F., Zhao, R., & Yang, Y. (2011). Effects of online not taking formats and self-monitoring prompts on learning from online text: Using technology to enhance self-regulated learning. *Contemporary Educational Psychology*, *36* (4), 313-322.

Kramarski, B. & Michalsky, T. (2010). Preparing preservice teachers for self-regulated learning in the context of technological pedagogical content knowledge. *Learning and Instruction*, *20* (5), 434-447.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Lipsey, M., Puzio, K., Yun, C., Hebert, M., Steinka-Fry, K., Cole, M., Roberts, M., Anthony, K. and Busick, M. (2012) Translating the statistical representation of the effects of education interventions into more readily interpretable forms, Washington DC: Institute of Education Sciences

Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of Documentation*, *61* (6), 700-712.

Mayes, D. K., Sims, V. K., & Koonce, J. M. (2001). Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics*, 28, 367-378.

Morineau, T., Blanche, C., Tobin, L., & Gueguen, N. (2005). The emergence of the contextual role of the e-book in cognitive processes through an ecological and functional analysis. *International   Journal of Human-Computer Studies*, *62* (3), 329-348.

Muter, P., Latremouille, S. A., Treurniet, W. C. & Beam, P. (1982). Extended reading of continuous text on television screens. *Human Factors*, *24* (5), 501-508.

Muter, P. & Maurutto, P. (1991). Reading and skimming from computer screens and books: the paperless office revisited? *Behavior and Information Technology*, *10* (4), 257-266.

Narciss, S., Proske, A., & Koerndle, H. (2007). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior*, *23* (3), 1126-1144.

Nelson, T. O. & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: MIT Press.

Noyes, J. M. & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, *51* (9), 1352-1375.

Noyes, J. M., Garland, K. J., & Robbins, E. L. (2004). Paper-based versus computer-based assessment: Is workload another test mode effect? *British Journal of Educational Technology*, *35* (1), 111-113.

Oborne, D. J. & Holton, D. (1988). Reading from screen versus paper: there is no difference. *International Journal of Man-Machine Studies*, *28* (1), 1-9.

Oliver, R. (1994). Proof-reading on paper and screens: The influence of practice and experience on performance. *Journal of Computer-Based Instruction*, *20* (4), 118-124.

Online Learning Consortium. (2012). Changing Course: Ten Years of Tracking Online Education in the United States. *Online Learning Consortium*. Retrieved from: http://onlinelearningconsortium.org/survey_report/changing-course-ten-years-tracking-online- education-united-states/

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, *38* (1), 1-4.

Pew Research. (2012, September 27). *In Changing News Landscape, Even Television is Vulnerable*. Retrieved from: http://www.people-press.org/2012/09/27/in-changing-news-landscape-even-television-is-vulnerable/

Pintrich, R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, *92* (3), 544-555.

Pressley, M. (1995). More about the development of self-regulation: complex, long-term, and thoroughly social. *Educational Psychologist*, *30*, 207-212.

Pressley, M., Levin, J. R., Ghatala, E. S., & Ahmad, M. (1987). Test monitoring in young grade-school children. *Journal of Experimental Child Psychology*, *44*, 96-111.

Rothkopf, E. Z. (1971). Incidental memory for location of information in text. *Journal of Verbal Learning and Verbal Behavior*, *10* (6), 608-613.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: metacognition as part of a broader perspective on learning. *Research in Science Education*, *36* (1-2), 111-139.

Schuetze, C. (2011). *Textbooks Finally Take a Big Leap to Digital*. Retrieved from http://www.nytimes.com/2011/11/24/world/americas/schoolwork-gets-swept-up-in-rush-to-go- digital.html?pagewanted=all

Shaikh, A. D. (2004). Paper or Pixels: What are People Reading Online? *Usability News*, *6* (2).

Sitzmann, T., Bell, B. S., Kraiger, K., & Kanar, A. M. (2009). A multilevel analysis of the effect of prompting self-regulation in technology-delivered instruction. *Personnel Psychology*, *62* (4), 697-734.

Spencer, C. (2006). Research on Learners' Preferences for Reading From a Printed Text or From a Computer Screen. *Journal of Distance Education*, *21* (1), 33-50.

The Council for Research Excellence. (2009, April 3). *A Day in the Media Life – Some Findings From   The Video Consumer Mapping Study*. Retrieved from http://www.researchexcellence.com/ CREVCMNCTACableShow_040309.pdf

The Sloan Consortium (2013, January 8). *Changing Course: Ten Years of Tracking Online Education in the United States*. Retrieved from: http://sloanconsortium.org/publications/ survey/changing_course_2012

Van den Boom, G., Paas, F., van Merrienboer, J., & van Gog, T. (2004). Reflection prompts and tutor feedback in a web-based learning environment: effects on students' self-regulated learning competence. *Computers in Human Behavior*, *20* (4), 551-567.

Wastlund, E. (2007). *Experimental studies of human-computer interaction: Working memory and mental workload in complex cognition*. (Doctoral dissertation). Retrieved from https://gupea.ub.gu.se/bitstream /2077/4693/1/gupea_2077_4693_1.pdf

Wastlund, E., Reinikka, H., Norlander, T., & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in Human Behavior*, *21*, 377-394.

Wilkinson, R. T. & Robinshaw, H. M. (1987). Proof-reading: VDU and paper text compared for speed, accuracy and fatigue. *Behaviour and Information Technology*, *6* (2), 125-133.

Wippich, W. (1981). Does a purchasing situation improve the prediction of one's own memory performance in preschool-aged children? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *13* (4), 280-290.

Wu, M. & Chen, S. (2011) Graduate students' usage of and attitudes towards e-books: experiences from Taiwan. *Program*, *45* (3), 294-307.

Ziefle, M. (1998). Effects of display resolution on visual performance. *Human Factors*, *40* (4), 554–568.

Zimmerman, B. & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, *223*, 614-628.

Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist*, *30* (4), 217-221.

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: an analysis

of exemplary instructional models. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Self-*

regulated learning: from teaching to self-reflective practice (pp. 1-19). New York:

Guilford Press.