Summer 2018

# Using instructed response times to compare inattentive responding across paper and online modalities: measuring research participant inattention

Olivia B. Brooks
*Eastern Washington University*

# USING INSTRUCTED RESPONSE ITEMS TO COMPARE INATTENTIVE RESPONDING ACROSS PAPER AND ONLINE MODALITIES: MEASURING RESEARCH PARTICIPANT INATTENTION

---

A Thesis

Presented to

Eastern Washington University

Cheney, Washington

---

In Partial Fulfillment of the Requirements

for the Degree

Master of Science in Clinical Psychology

---

By

Olivia B. Brooks

Summer 2018

THESIS OF OLIVIA BROOKS APPROVED BY

_____DATE_____
AMANI EL-ALAYLI, GRADUATE STUDY COMMITTEE

_____DATE_____
PHILIP WATKINS, GRADUATE STUDY COMMITTEE

_____DATE_____
MERRI HARTSE, GRADUATE STUDY COMMITTEE

# Using Instructed Response Items to Compare Inattentive Responding across Paper and Online Modalities: Measuring Research Participant Inattention

Olivia Brooks

Eastern Washington University

Abstract

It is important to examine the equivalence of paper and online data collection methods across several domains. The current study compared paper versus online modalities from a data quality standpoint, with a specific focus on inattentive (i.e., careless) responding by using an easily-implemented method to capture participants' careless responses: interspersing instructed response items (e.g., "Please select option 3, 'Strongly disagree,'") throughout a collection of 15 established measures. A between-subjects design compared the percentage of instructed response items missed (hence, inattentive responses) across three conditions: 1) lab paper, 2) lab online, and 3) non-lab online. The non-lab online condition was predicted to have the most inattentive responses, followed by the lab online condition, and lastly the paper condition. This hypothesis was based on the tendency for distracting and multitasking activities to increase in uncontrolled environments (Liu, 2012; Moisala et al., 2016). Additionally, sustained attention decreases with online displays and people tend to spend less time on in-depth reading (Birkerts, 1996) due to fewer spatial cues necessary for comprehension (Liu, 2012; Mangen et al., 2013). 160 college students participated in the study. The results supported my hypothesis that participants in the non-lab online condition would be most inattentive as the non-lab condition missed significantly more instructed response items than the two lab conditions. However, participants in the lab online condition were no less attentive than those in the lab paper condition. Ultimately, these results indicate online data collection is a viable option in lab studies which require sustained attention and that implementing a simple intervention such as instructed response items is an effective and necessary step in identifying inattention. Online data collection in non-lab environments

may be developing in a promising direction, but further research is needed to identify and counter the increased tendency for inattention in those environments.

# Using Instructed Response Items to Compare Inattentive Responding across Paper and Online Modalities: Measuring Research Participant Inattention

Online data collection is becoming a popular choice for researchers, with benefits to numerous aspects of scientific studies, ranging from population access to protocol and statistical design. For example, online data collection increases convenience, anonymity (i.e., privacy), and accessibility to larger and more nontraditional populations less likely to participate in paper-and-pencil studies, such as adolescent, international, or rural populations (Dias, Maroco, & Campos, 2015; Duffy, Smith, Terhanian, & Bremer, 2015). It also reduces paper waste, an important factor in current sustainability practices. Many researchers also praise online data collection for direct input of participant data, as opposed to first translating data from paper to computer databases, which leaves data more vulnerable to human errors like duplication and loss (Dias et al., 2015).

The appeal of these and additional benefits has steered many researchers away from the traditional, lab paper format of data collection, particularly in university settings such as that of the current study. Yet, before abandoning paper for online data collection altogether, we must first evaluate the equivalence of the two modalities. Changing the modality and display of a measure may inherently affect its psychometric qualities, thus it is important to compare the equivalence of paper and online methods across several domains (Dias et al., 2015). These considerations are not new; however, previous research has yielded inconsistent findings (Weigold, A., Weigold, I., & Russell, 2013). The current study advances the comparison of paper and online modalities from a data quality standpoint, with a unique focus on the issue of inattentive responding.

In psychological research, we are typically interested in examining variables that may lead to low quality data, as we want to control for "bad" data that could impair the integrity of our experimental studies. To control for potential data quality detriments, researchers strive to identify and implement appropriate methods that avoid common biases, such as response bias, and choose relevant statistical analyses. One variable that particularly leads to low quality data is inattentive (i.e., careless) responses from study participants. The current study identifies inattentive responses through use of a simple tool, instructed-response items (e.g. "Please select option 3, 'Strongly disagree,' for data checking purposes."), to explore data quality across paper and online modalities. It is hoped that the present study will expand upon other data quality research focused on response behavior and will steer future university researchers (or otherwise) towards the most optimal data collection method. Given that universities have largely embraced the use of online study formats, and college students are a common data collection source in psychological research (Casler, Bickel, & Hackett, 2013; Gordon, Slade, & Schmitt, 1986), the current study focuses on data gathered from a university sample.

**History and Definition of Inattentive Responding**

In any form of data collection, there are bound to be responses from participants that do not truly represent the targeted measure (Meade & Craig, 2012). Though the minority, these responses are problematic because they falsely influence the results of a study. The history of problematic responding in research has highlighted the importance of designing study protocols with participant behavior in mind. One may recall some memorable interventions designed to combat response issues; for example, the developers of the Minnesota Multiphasic Personality Inventory (MMPI), a hallmark

personality assessment in the field of psychology first published in 1942, were pioneers in early detection of and accounting for responding issues (Butcher, 1989). The MMPI-2 (revised) incorporates validity scales to capture certain participant response behaviors. For example, the L [Lie] Scale detects unsophisticated and naïve attempts of participants to present themselves in a favorable way, and the K [Korrection] Scale measures clinical defensiveness (Rodgers, 2008). There are also two scales which measure response inconsistency. Specifically, VRIN [Variable Response Inconsistency] is revealed when two similar items receive very different responses (e.g., True for "My father was a good man" and False for "My father was nice.") and TRIN [True Response Inconsistency] is shown when two opposite items receive similar responses (e.g., True for both "My father was a good man" and "My father was mean"). In fact, if any of these validity scales are clinically significant, the accuracy of the entire assessment must be questioned. The development of the MMPI validity scales was one important model that shows how psychometric design can facilitate the identification and resolution of participant response issues to ensure data integrity.

Unlike the MMPI-2, the present study focuses on just one form of problematic response to evaluate data quality, specifically, inattentive or careless responding. Not paying attention to study items is an important factor that can cause invalid responses from participants. Nichols, Greene and Schmolck (1989) notably suggest the existence of two types of troublesome responses: 1) content responsive faking, and 2) content nonresponsivity. Content responsive faking means that the content of a measure's item influences the participant's response, and that the response is not entirely accurate (Nichols et al., 1989). Said differently, the participant responds in a socially-desirable

way based on what he or she judges to be the intent of the item. With this type of faking, the participant does not have any problem attending to measure items; the problem is the catered response. Alternatively, content nonresponsivity means an item's content does not have any bearing on the response because the participant is not paying attention to the item's content in the first place. Content nonresponsivity is the type of inaccurate response that the current study seeks to measure to determine equivalence of online and paper modalities.

An important distinction to note is that in previous literature, content nonresponsivity has often been referred to as "random" responding. At face value, this description appears accurate because we think of "random" as denoting an unsystematic pattern; a choice without conscious decision (Beach, 1989; Berry et al., 1992). However, Meade and Craig (2012) suggest that the terms "inattentive" or "careless" responding are preferred over the term "random" responding. This is because it is possible for participants to be purposefully "random" (i.e., systematic) in their response pattern, which is not the same as content nonresponsivity. Thus, the terms "inattentive" and "careless" responding are more valid and precise descriptions of the unsystematic behavior that occurs in study participants. "Inattentive" responding and "careless" responding are used interchangeably in the current study.

As the aim of the present research is to capture inattentive responses, a procedure for measuring inattention must be implemented. This was accomplished by embedding instructed-response items designed to detect inattentive responding throughout the study. An example is "Please select option 3, 'Strongly disagree,' for data checking purposes." Variations of instructed response items have appeared in past research, such as

Oppenheimer's (2009) "instructional manipulation checks" which required participants to write, "I read the instructions" at the beginning of the study; though none of these studies align with the aims of the current research because they did not target inattentive behavior throughout the study. In Oppenheimer's study the instructions were only presented once at the beginning of the study to verify participants read the prompt, before participants had even provided any responses to the study. Casler et al. (2013) embedded a textbox in which participants were instructed to write one thing that they learned from the previous task. The differences with those styles of instructed response items is they were not concealed within individual study items, they requested participants to self-report their performance, and ultimately, they were used to assess participants' understanding of the study's content. This differs from the current instructed response items which are by design benign items embedded within the study to unobtrusively detect whether participants are responding carelessly. The mechanism for identifying inattentive responding via instructed response items is that a careless responder would not attend to the instructions and therefore would respond incorrectly or completely miss the item, cueing the researcher of inattentive behavior. Instructed response items are an easily-implemented and proactive approach to measuring low-quality data due to inattentive responding.

Instructed response items provide a unique opportunity to indirectly capture trends that are normally outside of researcher control, such as behavioral fluctuations missed with conventional indicators of response issues. Conventional methods include conducting consistency checks and flagging or eliminating "don't know," incomplete, or answers with patterns, like straight-lining (choosing answers in a straight line down the

page; Greszki, Meyer, & Schoen, 2015). While generally helpful, these methods only detect low quality data answers that fit well-known patterns (Greszki et al., 2015). Another common strategy is to use time (i.e., response rates) as an indicator of low quality data. Like the current study, Grezski et al. (2015) were concerned about the problems that online data collection raises and how to detect more sophisticated "bad" responses. The authors utilized time (i.e., response rates) to investigate response behavior by analyzing case-wise (average rate over the entire study) and page-wise (average rate of each page of the study) response rates. Based on the response process model (Tourangeau, 1984), Greszki et al. posited that very short response times indicate low quality data, which stems from inattention on the part of study participants. It was revealed that quick responses were not necessarily associated with inattention. Also, participants' attention to items varied during the study, and attention even changed within each page, thus it is important to develop tools to detect those changes. Whereas calculating response time may be a useful strategy to get an overview of respondent activity, instructed response items are a viable method of checking for such fluctuations in participants' attention because the researcher can embed them throughout the study.

Instructed response items may also be an optimal choice when study measures are not designed with their own sophisticated validity and reliability scales for content-based responding issues, such as VRIN and TRIN in the aforementioned MMPI-2. The current study focuses on a different type of problematic responding (inattention) because people might really have conflicting opinions that unintentionally get categorized as invalid with the VRIN and TRIN. Instructed response items have their own merit as benign statements that are easily embedded within a survey. In other words, they are not meant to elicit any

reaction from the participants other than to follow the instruction, and the procedure for implementing them is replicable with most types of data collection.

Instructed response items also avoid biases and inaccuracies that occur when relying only on self-reported attentiveness. For example, in Woolliscroft, TenHaken, Smith, and Calhoun (1993), academically lower-performing students tended to rate their performances higher than did their peers on an initial self-assessment, while higher-performing students rated themselves lower. The authors also found weak correlations between students' final self-assessment ratings and the ratings given by faculty. Like many other studies, Woolliscroft et al. concluded that there is poor agreement between self-assessments of performance and external measures of performance. Oppenheimer's (2009) study also found evidence to support the inaccuracy of self-reported performance-related behavior, such that participants who failed the instructional manipulation check in the study (i.e., did not acknowledge that they read the study instructions) reported statistically the same level of motivation to complete the study as those who passed. If self-reported performance were accurate, then participants who missed the instructional manipulation check should have reported lower motivation to complete the study and vice versa. These findings apply to the present study as we want to use the most accurate method to identify careless respondents as possible, while using caution to avoid relying on just one method. The implicit drawbacks to self-evaluation of performance in previous research are one reason why the present study used instructed response items to identify careless responses, and not just self-reported performance.

**Previous Research on Paper versus Online Modalities**

Past interest in online and paper comparisons has primarily focused on determining the two modalities' psychometric equivalence; in other words, determining if and how electronic and paper formats are equal via various theoretical and technical approaches. However, the current study is the first of its kind to make this determination by measuring the specific response behavior of inattention, and, further, using instructed response items to do so.

A primary approach has been to study how administration method affects equivalency of already-established instruments. It is through the development and validation of instruments that researchers are able to quantify subjective/psychological phenomena with objective/physical phenomena (Scripture, 1983). Changing the modality and display of an instrument necessitates scientific study because doing so may inherently affect its psychometric quality (Dias et al., 2015). Even a seemingly insignificant difference could affect the integrity of the instrument. For example, is selecting answer choices by circling them on paper the same as clicking answer choices on a webpage? One might speculate that the answer to this question depends on a great number of factors, such as participant attention level, aesthetic continuity (i.e., does the online measure look the same), or testing environment. Because of the tendency for such extraneous variables to influence responses, it is important to compare different methods of administering instruments.

Several studies using established instruments have found no significant difference in overall psychometric equivalence between online and paper versions, though results have been mixed depending on the theoretical construct measured. One recent example of

a study that compared an established instrument to its online version was conducted by Dias et al. (2015) with the Weight Concerns Scale (Portuguese version; WCS). The WCS is a one-dimensional, five-item measure which utilizes a seven-point Likert scale for responses. Dias et al. compared the WCS among 100 college students who were randomly assigned to first complete an online or paper format. Then, after a one-week washout period, the students completed the other format. A link to the online version was emailed to participants which implies they may have completed the study in a location of convenience and not in a lab setting. The authors measured equivalence by analyzing psychometric sensitivity, construct validity, factorial invariance, concurrent validity, and reliability. Their results confirmed the psychometric equivalence of the online and paper Weight Concerns Scale. Specifically, appropriate fit of the factor structure in both modalities was observed and the various forms of validity and reliability were found to be adequate (Dias et al., 2015). The equivalence of the paper Weight Concerns Scale and its online version is a hopeful sign that many variables would be equal across paper and online conditions. However, the authors did not assess all forms of equivalence, such as inattentive responding. Moreover, the short duration of their study (only 5 items) likely required minimal effort from the participants to pay attention to all items. Had it been a longer study, there may have been more careless responding in one modality than the other because over time additional variables influence responding behavior. For example, studies such as Mangen, Walgermo, and Bronnick (2013) and Eden and Eshet-Akkalai (2012) caution researchers that short measures such as the WCS may not result in significant differences because their brief nature does not require an increased cognitive load. Indeed, Alexander and Singer's (2017) review found that students' comprehension

is only impacted by digital and print materials after reading materials exceed one page in length at minimum. These factors suggest a need for further exploration into paper and online modalities to verify their equivalence. The present study compares the modalities for a study with a much longer duration.

Other researchers have examined equivalence by determining whether populations that complete studies online are more prone to certain characteristics than those that participate in traditional in-person study administrations. It is important to note that these suspected problems exclusively relate to open online recruitment (i.e., self-selection to online studies). Suspected problems with online study populations broadly fall into psychological, attitudinal, behavioral, and demographic categories.

One suspicion that past researchers had regarding psychological differences is that online users may have higher levels of depression or maladjustment than traditional samples, especially back when online studies were first getting introduced as a research tool. This idea was refuted by several subsequent studies, however (Gosling, Vazire, Srivastava, & John, 2004). Other studies postulated that online users are more/less invested or motivated than traditional in-person samples (Casler et al., 2013). This has historically been a common concern in survey research that has evolved over time with the shift in technology. For example, this concern was relevant with the popularization of mailing in surveys versus conducting them person. However, little evidence has been found to support that level of investment significantly differs in online populations. (Gosling et al., 2004). Further, as we learned with Oppenheimer (2009), self-reported motivation is a poor predictor of performance anyway.

Another area of interest in the history of paper and online comparisons is the hypothesis that online study takers are more likely to be comprised of a younger demographic because of younger generations' familiarity and confidence in using technology. To investigate equivalence of demographics, Casler et al., (2013) compared participants recruited from a university, social media, and Amazon's Mechanical Turk (MTurk). MTurk is an international marketplace where employers post tasks that are better designed for human rather than computer labor (Casler et al., 2013). Researchers have found this to be an innovative platform for recruiting study participants, and MTurk samples are known to be more diverse than most college samples (Casler et al., 2013). Casler et al. gave the three samples the same behavioral test to complete. Upon review of the demographic information of the three samples, the authors observed differences such that the MTurk participants were more socioeconomically and ethnically diverse. However, when it came to the results of the behavioral test, there were no significant performance differences between the samples. The findings of Casler et al. suggest that though there may be demographic variations within multiple recruitment samples, online recruitment and testing can be as valid (or more) than traditional face-to-face data collection.

Ultimately, many of these suspicions have been refuted or have received mixed support in the literature, further reinforcing the need for new studies about online and paper modalities. In consideration of the problems that may occur with open online recruitment, the present study drew its participant pool from the same university population. Also, to avoid any sampling issues with self-selection to preferred study modality, the present study took care to assign participants to online and paper

conditions. Specifically, all participants initially believed that they were signing up for a lab study, but some were later assigned to complete the study online instead.

## I**mportance of Current Study**

To date, no study has compared online versus paper administrations on response behavior solely using instructed response items as indicators of inattention. It is important to understand data quality issues when comparing paper and online modalities for the several reasons reported here. As well-intentioned as a research project may be, low-quality data lead to inaccurate findings or low statistical power, which, in turn, may alter the conclusions made in science. Innumerable variables impact data quality, so the question becomes, how can we best identify and control for these variables?

At the level of individual studies, identifying inattentive responses will facilitate a clearer picture of the prevalence of inattentive responding, cleaner data sets, and more accurate data analyses (Meade & Craig, 2012). For example, standard statistical procedures like outlier analysis may not be comprehensive in capturing low-quality data. Consider the theoretical assumptions behind outlier analysis. The assumptions are that problematic responses will differ from the majority of responses in the data set, and that the occurrence of problematic responses is rare (Meade & Craig, 2012). However, if a researcher were to only use outlier criteria as the basis for "bad" data, careless participants may go undetected if they are responding in a manner that is in range with the rest of the data set, such as answering "slightly agree" for all items (Meade & Craig, 2012). It is reasonable to suspect that a portion of careless responses falls within expected range and may go undetected. For this reason, part of the aim of the current study is to discover the rate of inattentive responding and help establish prevalence in the literature.

Further, the existence of outliers first depends on the comparison of all data within the set. If any of the data are compromised by inattentiveness, the integrity of the data set and any subsequent analyses (outlier or otherwise) is damaged, leading to a slew of inaccuracies in interpretation. Combining standard statistical procedures with additional measures of inattention may be a more effective way to account for these problematic responses in the future.

The present study takes the investigation of low-quality data to a new level by using instructed response items to compare inattentive responding in online versus paper formats, for which no research currently exists. A comparison of data quality between online and paper formats is especially important because there is an inherent risk for more problematic responses due to the uncontrolled environment and total anonymity in online responding. With increased utility of online studies as a trusted and convenient source of data collection, it is imperative to determine if any differences exist in this modality compared to paper.

**Prevalence of Inattentive Responding**

Existing prevalence reports of inattentive, or careless, responding in both paper and online modalities are inconsistent (Meade & Craig, 2012). For online studies, this is partially attributed to limited research on inattentive responding specifically online, and therefore a limited opportunity to establish a pattern of prevalence. Moreover, not all data quality researchers craft their study design using indicators of inattentive responding, and among those who do, there is still variability when it comes to how inattention is measured. Studies have used a wide array of methods to measure inattention, including comparisons of response rates, self-reported levels of attention, monitoring brain activity,

consistency indices, and more (Meade & Craig, 2012). The present research hopes to contribute to the picture of prevalence of inattentive responding using an indirect, but simple measurement of instructed response items.

As technology is ever-changing, it is also important to investigate current issues related to online versus paper modalities because some of the problems present when computers first became popular may not be relevant anymore. For example, studies made on computers from 1985 cannot be compared with today's digital screens (Myrberg, & Wiberg, 2015). Not only have the devices evolved, but the social climate has too. The majority of college students may not have owned a computer several years ago, whereas today a great deal of people have regular access to computers. According to the U.S. census, in 2013, 83.8 % of U.S. households reported computer ownership, with 78.5 % of all household having a desktop or laptop computer, and 63.6 % having a handheld computer (phone or tablet). Thus, the current study will also add to the overall understanding of issues with research that uses technology which is now readily available.

**Potential Reasons for Increased Inattentive Responding in Online Administration**

The prediction in the present study was that online administrations of the study would result in more inattentive responses than a paper administration. This prediction relates to common sense principles of distractible environments, multitasking, and the differences in comprehension with physical versus digital text. Ideally, participant responses have four steps: 1) Reading the entire question (i.e., comprehension), 2) Retrieving relevant information from memory, 3) Forming a judgment, and 4) Responding by writing an answer or clicking on the computer (Greszki, Meyer, &

Schoen, 2015). All four steps must be complete for response data to be valid (Greszki et al., 2015). Inattentive responses come from an interruption in completing any of these steps, and sometimes inattentive responders are not even successful in completing the first step.

One cause of unsuccessful completion of the steps is distractibility. Participants who take the study at their convenience are in an uncontrolled environment, and with that come limitless opportunities for distraction from the task at hand. Common distractors include loud noises (e.g. music, talking), problems with the technology used to complete the study, increased number of stimuli in environment, and multitasking—such as eating and drinking while completing the study. In a lab setting, the environment is controlled, and distractions are kept to a minimum by ensuring no talking, a clean room, personal space, reliable computers, etc.

Similarly, individuals completing tasks in their chosen environments might be more likely to engage in multitasking. The current young adult generation is particularly prone to engage in multitasking associated with distraction compared to older generations (Moisala et al., 2016). Multitasking has become a larger part of everyday life because of the inundation of various forms of technology that are readily accessible. Younger generations have grown up with more exposure to technology and engage in multitasking behaviors daily, such as texting, browsing the internet, reading online articles, etc.

Moisala et al. (2016) were concerned with previous findings that extensive media multitasking decreases attentional control, so they conducted a study in which 149 adolescents and young adults (age 13-24 years) performed speech-listening and reading tasks. These tasks required participants to sustain attention despite distracting stimuli.

Participants' brain activity during tasks was monitored by functional magnetic resonance imaging (fMRI), and participants also self-reported their amount of daily media multitasking. Moisala et al. made a couple of discoveries. First, higher amounts of reported everyday media multitasking was associated with more errors when participants performed a sentence congruency judgment task with a distractor stimulus present. This aligns with other studies that demonstrate multitaskers are more distractible. Second, during distracted attention/decreased performance there was increased activity in the right prefrontal cortex, an area of the brain related to attentional control. These findings are important because they demonstrate that multitasking and distractibility are positively linked, and the association is actually visible in brain activity changes.

Other studies (e.g. Liu, 2012) have even found that people who engage in multitasking behaviors, like alternating between two tasks, may spend 50% more time completing those tasks compared to if they just concentrated on one task and then the other. Because of these reasons, it is assumed that non-lab online administration of studies would be even more vulnerable to inattentive responding than would lab online administration. The former would offer environments that could elicit more distraction and multitasking, setting the stage for more careless responses.

Another factor to consider when predicting which modality will have more careless responses is that the physical nature of paper text and digital text (i.e., computers and other electronic media) is inherently different. In other words, paper surveys are tangible, but online surveys are only tangible insofar as the reader's interaction with the device through which the survey is administered. Physical differences between computers and paper are important to investigate because they may impact participants' ability to

read and attend to the study's text. For example, characteristics of a computer screen like refresh rate, high contrast, and fluctuating intensity of light can all interfere with cognitive processing for long-term memory (Noyes & Garland, 2003). Studies have even shown that there are health risks from looking at computer screens, such as screen-related sleeplessness from the blue light disrupting melatonin production, and computer vision syndrome, a temporary condition with symptoms like headache, fatigue, and strained and dry eyes (Barthakur, 2013; Myrberg & Wiberg, 2015; Wood, Rea, Plitnick, & Figueiro, 2013).

To understand the cognitive implications of text on screens versus paper, researchers began with the basic steps of how humans interact with texts (scientific studies or otherwise). Chatfield (2015) suggested that we should imagine how the brain interprets texts as a physical landscape, where every act of reading requires "identifying the special species of physical objects known as letters and words, using much the same neural circuits as we use to identify trees, cars, animals…" (para. 2). Because reading texts involves analysis of how the text is displayed on an object, studies have investigated how exactly spatial cues aid (or hinder) the learning process. Consistent with the growth of technology from approximately the 1970s through the 1990s, researchers discovered that digital navigation issues, such as scrolling and lack of linearity of pages, may interfere with the learning process (Mangen et al., 2013). By 2009, scrolling bars were on 91% of webpages (Sanchez, & Wiley, 2009). As Chatfield implies, readers often recall pieces of information based on where in the physical landscape of the text it appeared (e.g., upper-right corner of page 5). Scrolling on computers can hinder the reading process by imposing a spatial instability, which degrades the reader's mental

representation of the text, and in turn, comprehension of information, especially for readers that have lower working memory capacity (Mangen et al., 2013; Sanchez & Wiley, 2009). This is partially why longhand notes versus typed notes in class lend themselves to better conceptual understanding and retention (Chatfield, 2015). In addition to the clear spatial cues of paper, the relative slowness of writing longhand requires more mental effort and summarization and therefore more conceptual understanding, versus copying and pasting on a computer (Mueller & Oppenheimer, 2014). In sum, because of these differences with online displays, people tend to spend less time on in-depth reading, concentrated reading, and have decreased sustained attention on digital media compared to paper (Birkerts, 1996). In other words, people are more likely to skim information on computers (Liu, 2012). This has important implications for the current study as the participants who complete the online format may exhibit more inattention than the paper format.

Previous studies found that differences in performance between paper and online modalities lessened with shorter texts that required less sustained attention. The topic of the text may also influence the cognitive load, such as a scientific essay versus a straightforward questionnaire (Mangen et al., 2013). Therefore, the current study took care to ensure that spatial cues found on paper were mimicked in the online modality as closely as possible, and that the study's content was easily understood by the target audience. For example, navigational issues were avoided in the present study by implementing minimal scrolling. The study was spread over 41 pages—the same as on paper.

Interestingly, much of the differences in learning performance in past research was attributed to self-reported preference for or attitude towards computers (Kretzschmar et al., 2013). Liu (2012) was one researcher who investigated how people's view of electronic texts impacts performance. Liu likens the attachment people have to books to "traditional attachment." They like the page numbering (i.e., spatial cues), sense of ownership, and even the smell of books. This issue was also accounted for in the current study by including an item asking about participant preference for paper or online formats.

Ultimately, some differences (albeit minor) exist when comparing the physical nature of computer studies and paper studies. Because past research has indicated potential for decreased performance with longer texts on computers and that a substantial number of participants still prefer paper texts, this is another reason it is predicted that the online conditions will have more inattentive responses than the paper condition.

**The Current Study**

The current study compared data quality in online versus paper modalities by identifying inattentive responses across three different types of administration: lab paper; lab online; and non-lab online (i.e., at participants' convenience using their own device). It was hypothesized that the non-lab online condition would result in the most inattentive responses, followed by the lab online condition, and then the paper condition. Inattentive responding was assessed by instructed response items embedded within the pages of the study. Similarly, I also expected the same pattern to emerge when examining the effect of study condition on some supplemental dependent measures: the percentage of study items read carefully, the percentage of time spent on other tasks, and the extent to which

participants said they paid attention to the study. It may be that the uncontrolled nature of the non-lab condition is most inattentive due to combined potential for distractions and tendency for less in-depth reading on screens.

The study administered to participants was comprised of well-established measures of personality, motivations and attitudes. Some of the measures served no purpose other than to give participants surveys in which instructed response items could be embedded. Others were included both for this purpose and as supplemental measures that might relate to inattentive responding. For example, an established measure of absent-mindedness was included for possible use as a covariate. Absent-mindedness describes those lapses of attention and memory failures which result in minor inconveniences, such as losing time searching for missing keys or trying to remember a grocery list (Carriere, Cheyne, & Smilek, 2008). In the case of the current study, absent mindedness may contribute to careless responding. The remaining relevant measures are indicated throughout the method section. Thus, a secondary purpose of this study was to examine possible individual differences that may relate to inattentiveness, given that a lot of data were readily available. Identifying these variables could point to some individual differences that researchers may want to measure and control for in future research.

## Method

### Participants

Participants were college students recruited from psychology courses through Sona, Eastern Washington University's online system for research study sign-ups. Students participated for required or extra course credit. This research was approved by the Eastern Washington University Institutional Review Board (IRB).

There were 160 total participants of which three were dropped due to not finishing the study or violating instructions by using their phone to complete the study. The age range was from 18 to 49 years, with an average age of 22.03. There were 21.66% participants who identified as male, 77.07% who identified as female, and 1.27% who identified as other (transgender or gender fluid). Various ethnic groups were represented in the sample: White/Caucasian (67.52%), Hispanic/Latino (16.56%), Biracial (8.92%), Black/African American (2.55%), Asian/Pacific Islander (1.91%), Native American (1.27%), and Other (1.27%). Regarding class level, 15.92% were freshman, 17.20% were sophomores, 36.31% were juniors, 28.66% were seniors, and 1.91% were post-baccalaureate.

**Design**

The current research used a between-subjects design. The independent variable was the study administration modality. Participants were assigned to one of three study modalities: 1) lab paper, 2) lab online, 3) non-lab online. The dependent variable was the percentage of instructed response items missed/incorrect, which indicated the amount of inattentive responses.

**Procedure**

In the study, participants were identifiable only by their Sona ID number and were assigned to one of three conditions in a between-subjects design: (1) lab paper condition, (2) lab online condition and (3) non-lab online condition. To do this, prospective participants used Sona to sign up for a lab timeslot scheduled for a later date, after reading a blurb explaining that they will be contacted via email with further information regarding the study. The study sign-ups were tabulated (listed by Sona code number) and

every third person within each timeslot was assigned to the non-lab online condition. Participants in the lab conditions presented at their assigned time and the researchers alternated assignment between the paper and online lab conditions. Participants assigned to the non-lab online condition were emailed (anonymously via Sona) to give them a link for the online version of the study. After obtaining 60 participants in each lab condition, the remainder of participants who signed up were all assigned to the non-lab online condition until the end of the academic quarter.

Participants in the lab conditions came to the lab at assigned timeslots in groups of up to eight at a time and completed the study either on paper or on the computer (online). Participants in both online conditions completed the study via Qualtrics, a secure, computerized system designed for online study administration. Qualtrics is compatible with both Windows and Macintosh. Participants were instructed not to use cell phones due to notable variations in display and navigation.

The study was comprised of 15 well-established self-report measures, demographic questions, and supplemental questions**.** There were 22 total instructed response items inserted throughout the collection of measures, averaging 0.61 per page/screen or 5.45% of study items. Upon completion of the study, participants in the lab conditions were given a piece of paper debriefing them on the study. For those in the non-lab online condition, an online version of the same debriefing page appeared at the end of the study.

**Materials**

The study was 41 total pages/screens and included 15 well-established self-report measures, demographic questions, and supplemental questions. The main section of the

study was the collection of measures, which had 402 total items, was 36 pages/screens in length, and required approximately 45-55 minutes to complete. Demographic and supplemental sections were comprised of 28 items, totaling 5 pages in length and requiring approximately 5-10 minutes to complete. Thus, the entire study took an average of approximately 1 hour to complete. Instructed response items were not included in the demographic and supplemental items. Most study items were in Likert scale format. Some supplemental items which requested participants to give study observations/feedback were in multiple choice and open response formats.

**Instructed Response Items**. Instructed response items are those items which indicate inattentive responding behavior. These items instruct participants to select a certain answer choice. Correct selections should indicate that participants are paying attention to the study, while incorrect/missed items should indicate that participants are not paying attention. The location in which the instructed response item appeared on each page varied from one page/screen to the next. The instructed response items followed this format, with the phrasing and selection choice varying: "Please select option 3, Strongly disagree, for data checking purposes." (Refer to Appendix A for the complete list of instructed response items used in this study.) The dependent measure of the study was the percentage of instructed response items participants missed. The percentage was calculated by dividing the number of missed instructed response items by the number of instructed response items of which the participants were exposed. A percentage of instructed response items missed was used instead of a summation because some participants chose not to complete particular measures within the study and were therefore not exposed to all 22 instructed response items. For those reasons the

percentage of instructed response items missed ($M$ = 1.84, $SD$ = 7.78) is a more accurate representation of participant inattentiveness.

**Study Details.** The following describes the collection of well-established self-report measures presented to participants in chronological order. Some of these measures have a secondary purpose because they measure constructs that may relate to inattentive responding. Scoring information and descriptive statistics are included for those particular measures: Balanced Inventory of Desirable Responding, Frustration Discomfort Scale, Academic Motivation Scale, Attention-Related Cognitive Errors Scale, Boredom Proneness Scale, Mindful Attention Awareness Scale, Conscientiousness subscale of the Mini-International Personality Item Pool, and the Barratt Impulsiveness Scale.

The Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984) is comprised of 40 items designed to assess two dimensions of social desirability: impression management and self-deceptive enhancement (Leite & Beretvas, 2005). Impression management is defined as, "intentional faking of responses to create a socially-desirable image," (Leite & Beretvas, 2005, p. 144). Self-deceptive enhancement is, "the tendency to give honest but positively biased self-reports," (Leite & Beretvas, 2005, p. 144). Participants respond to statements using a scale of seven possible responses, ranging from 1, *not true*, to 7, *very true*. "*Once I've made up my mind, other people can seldom change my opinion*" is an example from the impression management scale (Paulhus, 1984). Despite the Marlowe-Crowne Social Desirability Scale (MCDS) being the most commonly used social desirability bias assessment, the BIDR was chosen over the MCSDS because of its ability to capture both impression management and self-

deception constructs whereas the MCSDS only measures one dimension, an individual's need for approval (Crowne, & Marlowe, 1960). To score the BIDR, certain items are reverse-scored and then all items are scored dichotomously by assigning the value of 1 to extreme answers (either 6 or 7) and the value of zero to the rest. The BIDR was one of the instruments the current study investigated as possibly being related to inattentive responding by performing a Pearson correlation. It was speculated that the impression management subscale ($M = 0.31$, $SD = 0.17$, $\alpha = .73$) may be negatively correlated with inattention such that participants who care more about image and approval might miss less instructed response items. Supplemental analysis was not conducted with the self-deception subscale as it did not seem like lying to oneself was relevant to inattentive responding.

Harrington's (2005) 28-item revised four-factor Frustration Discomfort Scale (FDS) measures low frustration tolerance through a series of statements to which participants respond with their level of belief on a scale from 1, *absent*, to 5, *very strong*. For example, "*I can't stand doing things that involve a lot of hassle*" (Harrington, 2005). This scale incorporates four factors: emotional intolerance, entitlement, discomfort intolerance, and achievement (Harrington, 2005). The emotional intolerance dimension involves intolerance of emotional distress; the entitlement dimension involves intolerance of unfairness and frustrated gratification; the discomfort intolerance dimension involves intolerance of difficulties and hassles; the achievement dimension involves intolerance of frustrated achievement goals (Harrington, 2005). To provide an overall assessment, all items were averaged together. Higher scores on the FDS indicate higher frustration discomfort (i.e., lower frustration tolerance). The FDS was another measure believed to

possibly be correlated with inattention, such that those with higher frustration discomfort in general may miss more instructed response items ($M$ = 2.89, $SD$ = 0.66, α = .93). It could be that participants who have a lower frustration tolerance would get frustrated during my lengthy study and their frustration would compete with their ability to fully attend to study items. Two of the subscales, achievement ($M$ = 3.14, $SD$ = 0.79, α = .77) and discomfort intolerance ($M$ = 2.36, $SD$ = 0.72, α = .79), were also included in supplemental analyses. Discomfort intolerance may be a relevant aspect of how overall frustration intolerance relates to inattentive responding because participants who do not tolerate difficulties or hassles may think of my study as difficult or a hassle, making them frustrated and in a state of lessened attention. Frustration pertaining to blocked achievement goals may also tie in with inattention because participants may get frustrated with obstacles or distractions while trying to achieve the goal of completing the study.

The revised form of the Gratitude Resentment and Appreciation Test (GRAT-R) is a 44-item measure which captures trait gratitude (Thomas & Watkins, 2003; Watkins, Woodward, Stone, & Kolts, 2003). Trait gratitude is a predisposition to experience the feeling (i.e., state) of gratitude (Watkins et al., 2003). The GRAT-R instructs respondents to indicate their level of agreement with each statement using a nine-point Likert scale, ranging from strongly disagree to strongly agree. For example, "*I think that it's important to pause often to 'count my blessings*." The revised version was selected over a short form because the current research wanted a lengthy study requiring sustained attention.

The Academic Motivation Scale College Version (AMS 28) is a 28-item measure that covers seven constructs of motivation relating to academics. These seven factors are intrinsic motivation (IM) to know, IM towards accomplishment, IM to experience

stimulation, extrinsic motivation (EM) identified, EM introjected, EM external regulation, and, lastly, amotivation (Vallerand et al., 1992). Participants use a 1-7 Likert scale (1, *do not agree at all*, to 7, *completely agree*) to indicate the extent to which 28 statements correspond to reasons why they attend college (e.g., "*Because my studies allow me to continue to learn about many things that interest me*"). Averages of each subscale must be computed to score this measure. Five subscales of the AMS were included in Pearson correlations due to potential relation to inattention. IM towards accomplishment ($M$ = 4.90, $SD$ = 1.55, α = .89) was chosen as it is about doing something for the satisfaction of accomplishment (Vallerand et al., 1992). EM identified ($M$ = 6.04, $SD$ = 0.96, α = .71) was chosen as it is about doing something because one has decided to do it, although it is not enjoyable (Vallerand et al., 1992). EM introjected ($M$ = 5.30, $SD$ = 1.51, α = .88) was chosen as it is about doing something because one pressures themselves to do it (Vallerand et al., 1992). EM external regulation ($M$ = 5.79, $SD$ = 1.28, α = .84) was chosen as it is about doing something because one is pressured by someone else to do it (Vallerand et al., 1992). Lastly, amotivation ($M$ = 1.73, $SD$ = 1.07, α = .82) was chosen as it is about having neither intrinsic or extrinsic motivation to do something (Vallerand et al., 1992). Those factors may have to do with how much or how little the participants' cared about the study items.

The 10-item Simplified 'Type A' Questionnaire (STAQ; Karlberg, Krakau, Sjödén, & Undén, 1997) is a short self-report measure that predicts Type A behavior. There are four factors: impatience, time pressure, hostility, and anger (Karlberg et al., 1997). Respondents indicate the extent to which they agree to 10 statements (e.g., "*I*

*never find enough time*") using a scale from 0 to 2 (0, *do not agree at all*, 1 *partly agree*, and 2, *fully agree*).

The 26-item Academic Entitlement Questionnaire (AEQ; Kopp, Zinn, Finney, & Jurich, 2011) measures five factors of academic entitlement and requires participants to respond to items on a scale from 1, *strongly disagree*, to 7, *strongly agree* (Kopp et al., 2011). For example, "*If I don't do well on a test, the professor should make tests easier or curve grades*."

The Attention-Related Cognitive Errors Scale (ARCES; Cheyne, Carriere, & Smilek, 2006) was chosen as one of two measures of absent mindedness. The ARCES measures the frequency with which one experiences a range of everyday cognitive failures which are most likely caused by lapses in attention. The ARCES is a 12-item questionnaire that uses a Likert scale of five possible responses, ranging from 1, *never*, to 5, *very often*. An example of an item is, "*I have lost track of a conversation because I zoned out when someone else was talking*" (Carriere et al., 2008). A higher score indicates a great number of slips. The ARCES was revised by its creators in 2008 to replace one item that was less related to the overall scale, and to slightly reword two items (Carriere et al., 2008). The revised version was used in the present study. Because the ARCES measures absent-mindedness it was included in Pearson correlations to explore if a greater number of cognitive slips in general was related to more inattentive responses in the present study ($M = 3.47$, $SD = 0.78$, $\alpha = .89$).

The Narcissistic Personality Inventory (NPI) was originally developed by Raskin and Terry in 1979 and further validated in 1988 and has become the most widely used narcissism inventory (Gentile et al., 2013). The instrument is a 40-item self-report

assessment of grandiose narcissism (Sherman et al., 2015). To align with the current study's aim to maintain a longer duration, the original 40-item instrument was used, rather than the NPI-13 (Gentile et al., 2013). The NPI provides two statements per item and instructs the participant to choose the statement that they most agree with. For example, for Item 1, participants must choose either A, "*I have a natural talent for influencing people*" or B, "*I am not good at influencing people*."

The 28-item Boredom Proneness Scale (BPS; Farmer & Sundberg, 1986) assesses one's proneness toward experiencing boredom. To complete the BPS, participants respond to statements using a Likert scale ranging from, 1 *strongly disagree*, to 7, *strongly agree*. Statements reflect situations in which people are likely to become bored, and related personal characteristics of boredom (Carriere et al., 2008). For example, "*It is easy to concentrate on my activities*" (reversed-item). The current study chose the 7-point Likert scale format of the BPS due to increased sensitivity when compared with the True/False format (LePera, 2011). Scoring the BPS requires reverse-scoring certain items and then averaging the scores. Higher scores indicate higher boredom proneness ($M$ = 3.59, $SD$ = 0.65, $\alpha$ = .80). The BPS measures constructs that may be related to inattention, thus it was included in correlation analyses to investigate if participants who tend to be bored were less attentive to instructed response items. Participants with higher boredom proneness may have been bored during my study and therefore less interested in paying close attention to the items.

The Subjective Happiness Scale (SHS; Lyubomirsky & Lepper, 1999) measures global subjective happiness. Participants respond to four items using a seven-point Likert scale, with higher scores reflecting greater happiness (Lyubomirsky & Lepper, 1999). As

an example, one item prompts participants to rank their happiness level in response to, "Compared to most of my peers, I consider myself …"

The Mindful Attention Awareness Scale (MAAS; Brown & Ryan, 2003) was the second measure of absent-mindedness. The 15-item MAAS assesses mindlessness in everyday situations (Brown & Ryan, 2003). This instrument employs a Likert scale that ranges from 1, *almost never*, to 6, *almost always* (Brown & Ryan, 2003; Carriere et al., 2008). For example, "*I find myself doing things without paying attention*" (Brown & Ryan, 2003). Higher scores on the MAAS indicate more absent-mindedness. This instrument was also included in supplemental analyses to examine its relationship with participant inattentiveness ($M = 3.58$, $SD = 0.78$, $\alpha = .74$). Given that participants who score higher on the MAAS are more absent-minded they might have more incidents where their attention is elsewhere, potentially causing them to miss more instructed response items.

The Big Three Perfectionism Scale (BTPS) is a newly-developed scale that measures dispositional perfectionism (Smith, Saklofske, Stoeber, & Sherry, 2016). It is a 45-item questionnaire that employs a five-point Likert scale ranging from 1, *strongly disagree*, to 5, *strongly agree* (Smith et al., 2016). Participants must respond to 45 statements using the 5-point scale, such as, "*Making even a small mistake would upset me*." The BTPS assesses three global factors composed of 10 lower-order core characteristics of perfectionism.

The Mini International Personality Item Pool (Mini-IPIP; Donnellan, Oswald, Baird, & Lucas, 2006) is a 20-item short form of the original International Personality Item Pool Five Factor Model (Goldberg, 1999). It measures the Big Five personality

factors, Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness, by having respondents report the extent to which they agree with 20 statements. For example, "*I am the life of the party*" (Donnellan, 2006). Options range from 1, *very inaccurate*, to 5, *very accurate* (International Personality Item Pool, 2017). The conscientiousness subscale items were averaged together ($M = 3.50$, $SD = 0.81$, $\alpha = .42$), and the measure was included in supplemental Pearson correlations due to speculation that participants who score higher on the conscientiousness scale might miss less instructed response items, therefore being overall more attentive. Conscientious people are careful and vigilant thus they may be inherently more attentive when completing my study.

The Satisfaction with Life Scale (SWLS) is a five-item questionnaire designed to measure global life satisfaction (Diener, Emmons, Larsen, & Griffin, 1985). Participants indicate the degree to which they agree with each statement by using a seven-point Likert scale that ranges from strongly disagree to strongly agree. For example, "*The conditions of my life are excellent*."

The Barratt Impulsiveness Scale is the most widely cited measure of impulsiveness (BIS; Patton, Stanford, & Barratt, 1995; Steinberg, Sharp, Stanford, & Tharp, 2013). Participants report the extent to which 30 statements regarding personal characteristics apply to them by using a 4-point scale (1, *rarely/never*, 2, *occasionally*, 3, *often*, and 4, *almost always/always*; McLeish & Oxoby, 2007). For example, "*I plan tasks carefully*." Certain items must be reversed before scoring, and then items for each subscale are averaged together. The three subscales of the BIS are attentional impulsiveness, which involves focus and thought control factors, motor impulsiveness,

which involves perseverance and unplanned action factors, and third, non-planning impulsiveness, which involves careful consideration of choices and problem factors. Pre-established sub-factors were drawn from the three subscales to investigate possible relation to inattention. They were attention ($M$ = 2.31, $SD$ = 0.56, α = .63) and perseverance ($M$ = 1.84, $SD$ = 0.50, α = .29). Participants' ability to concentrate or focus (attention), and their level of determination (perseverance) may be key components of impulsiveness that affects how much participants paid attention to the study. Participants with higher scores on either of these factors indicate greater impulsiveness, thus those participants may be inclined to greater carelessness when completing studies.

General demographic information about participants was collected using five items requesting age, race/ethnicity, gender, class level, and current academic course.

Supplemental questions were comprised of Likert scale items, free response, and multiple choice. (Refer to Appendix B for a complete list of the 23 supplemental questions.) In addition to data collection, the following items had a secondary purpose as they were investigated for correlations with inattentive responding: preference for electronic formats ($M$ = 2.75, $SD$ = 1.45; response scale 1, *strongly disagree*, to 7, *strongly agree*); values social science research ($M$ = 5.74, $SD$ = 1.26; response scale 1, *not at all*, to 7, *a great deal*); percent of study items read carefully ($M$ = 92.44, $SD$ = 10.07); percent of time spent on other tasks during study ($M$ = 5.17, $SD$ = 16.80); extent paid attention to study ($M$ = 5.99 , $SD$ = 1.14; response scale 1, *not at all*, to 7, *a great deal*; found the study boring ($M$ = 3.46, $SD$ = 1.71; response scale 1, *not at all boring*, to 7, *extremely boring*); and current GPA ($M$ = 3.37, $SD$ = 0.58). Specifically, participants who report a stronger preference for electronic formats over paper may miss more

instructed response items. It could be that a preference for electronic formats originates from more exposure to technology, and with that comes the propensity for multitasking. It was also speculated that participants who found the study more boring or spent more time on other tasks (i.e., multitasking) may miss more instructed response items because they lack motivation and/or get distracted. Conversely, higher participant GPA may negatively correlate with inattention as high-performing students might pay more attention to the study. Participants who reported a higher value of social sciences, a higher percentage of study items read carefully, or that they paid more attention to the study also may miss less instructed response items given their potentially greater motivation to complete the study.

In addition to correlation analyses, further exploratory analyses were carried out to examine the effect of study condition on some supplemental dependent measures: the percentage of items read carefully ($M$ = 92.44, $SD$ = 10.07), the percentage of time spent on other tasks ($M$ = 2.56, $SD$ =6.01, after excluding four large outliers), and the extent to which participants said they paid attention to the study, the latter of which was assessed on a scale from 1, *not at all*, to 7, *a great deal* (M = 5.99, SD = 1.14). It may be that the uncontrolled nature of the non-lab condition is most inattentive due to combined potential for distractions and tendency for less in-depth reading on screens.

## Results

*Tests of Hypotheses.* I used a One-Way ANOVA to examine the effect of study condition on percentage of instructed response items missed. The main effect of condition was significant, $F(2, 153) = 4.27$, $p = .016$, $\eta^2 = .05$. Tukey post-hoc tests were performed to compare the three conditions. A significantly greater percentage of

instructed response items was missed in the non-lab online condition than in both the lab paper condition ($p$ = .043) and the lab online condition ($p$ = .018), which did not differ from each other ($p$ = .925).

In the lab paper condition, participants missed approximately 1% of instructed response items ($M$ = 1.08, $SD$ = 2.96). In the lab online condition, participants missed approximately 0.5% of instructed response items ($M$ = 0.55, $SD$ = 1.72). In the non-lab online condition, participants missed an average of approximately 5% of instructed response items ($M$ = 4.90, $SD$ = 14.69).

These results supported my overall hypothesis that participants in the non-lab online condition would be the most inattentive (i.e., have the most instructed response items missed). My hypothesis that participants in the lab paper condition would outperform those in the lab online condition was not supported, as there was no significant difference between the two conditions.

*Individual Differences.* Since the current study was exploratory in nature, I also performed supplemental Pearson correlations between the percentage of instructed response items missed and measures of interest that may relate to inattentiveness. This could help reveal some individual differences that may contribute to inattentive responding in research studies. Refer to Table 1 for a complete list of correlations.

As shown in Table 1, some of the measures of interest related significantly to the dependent variable as might be predicted. Thus, both internal and external motivations may contribute to greater attentiveness during research studies. Amotivation in the Academic Motivation Scale College Version had the strongest relationship with percent instructed response items missed such that the more amotivated participants were, the

more inattentive they were, creating a positive correlation. The perseverance factor of the Barratt Impulsive Scale (BIS) was also positively related to percentage of instructed response items missed such that participants who scored higher on the perseverance scale (i.e., high in impulsiveness that pertains to perseverance) missed more instructed response items.

External motivational tendencies also related to inattention. The extrinsic motivation (EM) identified subscale of the AMS was related to inattention, but in this case, it was a negative correlation such that the more participants agreed that they do things because they decided to do them, but not because they are fun, the fewer instructed response items they missed. Also from the AMS, the EM external regulation subscale had a negative relationship with inattention such that the more participants agreed they do things because they are pressured by someone else to do them, the fewer instructed response items they missed. Further supporting the notion of external pressures reducing inattentive responding, a negative correlation was observed such that participants with higher impression management scores on the Balanced Inventory of Desirable Responding (BIDR) missed fewer instructed response items.

Some self-reported attitudes were also revealed to be significantly correlated with inattention. For example, participants who reported stronger preference of electronic formats missed more instructed response items. In addition, participants who reported greater value of social science research missed fewer instructed response items. Participants also seemed to have some awareness of their inattentiveness, given that participants who reported a higher percentage of study items read carefully missed fewer instructed response items (see Table 1).

Pearson analyses conducted between inattentive responding and the rest of the measures of interest did not result in significant correlations, contrary to expectations (see Table 1). This includes: overall frustration intolerance, discomfort intolerance, and achievement (Frustration Discomfort Scale; FDS); intrinsic motivation (IM) towards accomplishment and IM introjected (AMS); boredom proneness (Boredom Proneness Scale; BPS); conscientiousness (Mini International Personality Item Pool; Mini-IPIP); and attention (BIS). Inattention was also not significantly related with either measure of absent-mindedness, the Attention-Related Cognitive Errors Scale (ARCES) and the Mindful Attention Awareness Scale (MAAS).

Non-significant Pearson correlations with self-reported supplemental and demographic items are also listed in Table 1. Reported extent to which the participants' paid attention to the study, reported extent to which the participants found the study boring, reported percentage of time spent on other tasks, and GPA did not result in any significant correlations with inattentiveness.

*Supplemental Dependent Measures.* In addition to examining missed instructed response items, I examined possible condition effects on three supplemental dependent measures. First, I conducted a One-Way ANOVA to examine the effect of study condition on reported percentage of study items read carefully to further support my main finding that non-lab participants are the least attentive. The main effect of condition was significant, $F(2, 144) = 5.90$, $p = .003$, $\eta^2 = .08$. Tukey post-hoc tests were performed to compare the three conditions. Participants in the non-lab online condition reported that they read a significantly lower percentage of items with care compared to both the lab paper condition ($p = .004$) and the lab online condition ($p = .016$), which did not differ

from each other ($p$ = .816). Participants in the lab paper condition reported they read approximately 95% of study items carefully, ($M$ = 94.62, $SD$ = 6.93). In the lab online condition, participants reported they read approximately 93% of study items carefully ($M$ = 93.48, $SD$ = 8.62). In the non-lab online condition, participants reported they read approximately 88% of instructed response items ($M$ = 87.76, $SD$ = 13.94).

Similar results were found when conducting a One-Way ANOVA on the Likert item assessing the extent to which participants felt they attended to the items in the study. The main effect was significant, $F(2, 151) = 10.96$, $p < .001$, $\eta^2 = .12$. Tukey post-hoc tests revealed the mean in the non-lab online condition was significantly lower than that in either the lab online ($p$ = .006) or the lab paper condition ($p < .001$), which did not differ from each other significantly ($p$ = .255), indicating that participants reported the least attention in the non-lab environment. Referring to the Likert scale (1, *not at all*, to 7, *a great deal*), the average extent to which participants said they paid attention in the lab paper condition was 6.37 ($SD$ = 0.79), in the lab online condition was 6.05 ($SD$ = 1.08), and in the non-lab online condition was 5.36 ($SD$ = 1.37).

Though the Pearson analysis did not reveal a significant correlation between reported percentage of time spent on other tasks and the percentage of instructed response items missed (see Table 1), I still wanted to examine if there was an effect of study condition on reported time spent on other tasks, given that non-lab environments have more distracting stimuli that may make participants more susceptible to engaging in non-task-related activities. A One-Way ANOVA showed that the main effect of study condition was significant, $F(2, 142) = 5.90$, $p < .001$, $\eta^2 = .13$. Tukey post-hoc tests were performed to compare the three conditions. Participants in the non-lab online condition

reported they spent significantly more time on other tasks compared to both the lab paper condition ($p = .001$) and the lab online condition ($p < .001$), further solidifying that participants are more prone to distraction in uncontrolled environments. As with the other two dependent measures, the two lab conditions did not differ significantly from one another ($p = .556$). In the lab paper condition, participants reported that they spent 2% of the time on other tasks ($M = 1.94$, $SD = 3.93$). In the lab online condition, participants reported they spent 0.8% of the time on other tasks ($M = 0.82$, $SD = 2.22$). In the non-lab online condition, participants reported they spent 6% of the time on other tasks ($M = 6.32$, $SD = 10.05$).

**Table 1** *Correlations between Measures of Interest and Percent Instructed Response Items Missed*

| | **Percent Instructed Response Items Missed** | | |
|---|---|---|---|
| **Measure** | ***r*** | ***p*** | ***N*** |
| Balanced Inventory of Desirable Responding | ---- | ---- | ---- |
| Impression Management | -.18* | .028 | 156 |
| Frustration Discomfort Scale | .01 | .914 | 156 |
| Discomfort Intolerance | .14 | .087 | 159 |
| Achievement | -.03 | .681 | 159 |
| Academic Motivation Scale- College | ---- | ---- | ---- |
| Intrinsic Motivation (IM) towards Accomplishment | -.05 | .509 | 159 |
| Extrinsic Motivation (EM) Identified | -.27** | .001 | 159 |
| EM Introjected | -.09 | .242 | 159 |
| EM External Regulation | -.25** | .002 | 159 |
| Amotivation | .34** | <.001 | 159 |
| Attention-Related Cognitive Errors Scale | -.06 | .428 | 156 |
| Boredom Proneness Scale | .08 | .328 | 156 |
| Mindful Attention Awareness Scale | .02 | .813 | 156 |
| Mini International Personality Item Pool | ---- | ---- | ---- |
| Conscientiousness | -.06 | .452 | 155 |
| Barratt Impulsiveness Scale | ---- | ---- | ---- |

| | | | |
|---|---|---|---|
| Attention | .01 | .954 | 154 |
| Perseverance[a] | .32** | <.001 | 154 |
| Self-Report Supplemental Items | ---- | ---- | ---- |
| Preference of electronic formats over paper | .17* | .040 | 154 |
| Values social sciences | -.19* | .017 | 154 |
| Percent of items read carefully | -.25* | .002 | 147 |
| Percent time spent on other tasks [b] | .03 | .696 | 145 |
| Paid attention | -.13 | .105 | 154 |
| Study boring | .02 | .807 | 154 |
| GPA | -.14 | .098 | 148 |

[a]For perseverance, higher scores indicate higher impulsivity.

[b]After removal of four outliers on this measure.

*$p < .05$ **$p < .01$

## Discussion

The results of the current study inform us that there is indeed an effect of study administration modality on inattentive responding. Participants who completed the study outside of the lab on their own device missed significantly more instructed response items compared to the lab paper condition and compared to the lab online condition. Further supporting this notion, participants who completed the study online outside of the lab reported being less careful/attentive in their survey responding. Unexpectedly, the lab paper and lab online conditions were equal in the amount of instructed response items missed (as well as self-reported attentiveness). Therefore, my hypothesis that the non-lab online condition would result in the most inattentive responses was supported, but my hypothesis that the lab online condition would result in more inattentive responses than the lab paper condition was not supported.

Reflecting on the potential reasons for inattentive responding, my prediction that the online conditions would be more inattentive than the paper condition was based on

the online modality problems of increased distraction, greater multitasking, the lack of spatial cues necessary for comprehension (Alexander & Singer, 2017, Mangen et al., 2013; Sanchez & Wiley, 2009), and the tendency to spend less time on in-depth reading, concentrated reading, and sustained attention with digital text (Birkerts, 1996). My hypothesis that non-lab online participants would miss the most items out of all the conditions originated from the combined potential for all of those variables to occur when participants were in a non-lab setting using their own electronic device. I made the prediction that the lab online condition would miss more instructed response items than the lab paper condition because I thought even if the potential for distraction and/or multitasking decreased in both lab conditions, the lack of spatial cues and the tendency for less-concentrated reading in the online version of my study would still elicit more inattention than the paper version.

The study results lend several conclusions about inattention in online versus paper modalities. The main finding that participants were most inattentive outside of the lab aligns with previous studies (e.g., Mangen et al., 2013) which suggest there is a tendency to be more inattentive in natural (i.e., non-lab) settings due to the negative effects of distracting environments and multitasking on attentional control (Moisala et al., 2016). Participants in the current study had ample opportunity to engage in distracting and multitasking behaviors because the collection of measures was designed to be lengthy and requiring sustained attention. Indeed, participants in the non-lab online condition reported the most multitasking, suggesting that typical distractors like loud noises (e.g., music, talking), issues with the technology used to complete the study, eating/drinking, or using other websites or social media may increase and/or have a higher impact when the

environment is less structured. In fact, one of the supplemental items in my study prompted participants to describe any activities they engaged in during the study. Examples of their responses were: assisting children, watching television, listening to music, studying, using social media, reading emails, talking with friends, thinking about other things, eating/drinking, making phone calls, and texting. It is possible then that multitasking and distractions in the study environment influenced the percentage of instructed response items missed. It makes sense that participants in both lab conditions paid more attention to study items than participants in their chosen environment because lab settings are controlled, and distractions are kept to a minimum by ensuring no talking, no television, a clean room, personal space, reliable computers, etc. Another reason might be that participants feel more social pressure in the lab to take the study seriously, regardless of amount of distractions.

The lack of difference in inattentiveness between the lab paper and lab online conditions was unexpected but suggests that inattention may be more about study environment, and less about paper versus online modality. These results have beneficial implications. They suggest that despite inherent spatial differences across paper and online formats, efforts to translate one format to the other are succeeding and participants are able to maintain concentration. The present study ensured the online and paper versions were as identical in appearance as possible to facilitate equal conditions. For example, the computer pages had minimal scrolling and the number of pages/screens and layout of items were the same between conditions. This is valuable information for future research design because it suggests paper and online modalities are the same unless the study environment is uncontrolled. It is noteworthy, however, that more traditional

formatting in paper-and-pencil studies (versus the present version which appeared like the online version) might result in a different level of attentiveness and should thus be investigated in future research. Traditional paper-and-pencil formatting would be more familiar to participants, which may increase attentiveness.

The results from exploratory correlation analyses with measures of interest added to the general picture of inattention. Both internal and external motivational tendencies related to inattentive responding. For example, the more amotivated participants were and the greater impulsivity they had regarding perseverance, the more inattentive they were. Perhaps parallels can be drawn given that both constructs pertain to participants' general lack of tenacity.

Regarding external motivational tendencies, participants missed fewer instructed response items if they scored higher on measures which indicated they cared more about managing others' impression of them, tended to do things because they are pressured by someone else to do them, or tend to do things because of decision rather than pleasure. This applies to the current research as participants signed up for my study to earn class credit. It is likely participants took the study more seriously (hence, paid more attention) regardless of condition if they felt outside pressure to complete the study or accepted that it was not fun, but they decided to complete it anyway. The lengthy duration of the study and requirement of sustained attention may have also given the impression that the study should be taken seriously, explaining why participants especially high in impression management paid more attention. Thus, it is not just the study setting that may affect inattentive responding.

Motivation levels are also domain-specific, which could influence inattentiveness in specific domains. For example, participants who placed a higher value on social science research missed significantly fewer instructed response items in this study, which was conducted in the social science field. This result reinforces other findings from my study showing that participants who took it seriously or cared in some capacity were less inattentive. This suggests that if researchers can clearly emphasize the importance of their study before participants complete it, then problematic responding may be reduced, as most participants want to avoid "messing up" critical data collection. For example, researchers could insert a statement about the results being used for publication in a scientific journal. Future research should investigate the possible benefits of using these procedures.

Some measures of interest were not significantly correlated with inattention as expected. These included boredom proneness, how boring participants found the study to be, general intolerance of frustration, intolerance of difficulties or hassles, and intolerance of frustrated achievement goals. It is challenging to interpret why constructs such as these were not related to inattention when seemingly similar constructs like amotivation or perseverance were related. It could be that drives such as motivation and perseverance (or lack thereof) could override boredom and frustration, depending on the task. Specifically, participants earned course credit for completing my study, thus they may be more motivated to attend to the items despite level of boredom or frustration. It may also be that participants simply did not experience boredom or frustration while completing this particular study.

The two measures of absent-mindedness, the Attention-Related Cognitive Errors Scale (ARCES) and the Mindful Attention Awareness Scale (MAAS), also were not correlated with inattention, meaning that everyday memory slips or lapses in attention did not relate to the percentage of instructed response items missed. This implies that the type of inattention measured in the ARCES and MAAS may be different than the inattention captured with instructed response items in a study. For example, everyday lapses in attention may occur regardless of the stimulus as they are general incidents and not necessarily provoked by completing a study. Another possibility is that completing research studies is an easy task that does not require the level of focus that other everyday activities may require.

Given that previous literature has found poor correspondence between self-assessed performance and external measures of performance (Oppenheimer, 2009; Woolliscroft et al., 1993), it should not be surprising that there were no significant relationships between inattentive responding and self-reported attention paid to the study or percentage of time spent on other tasks during the study (i.e., multitasking). There was, however, a relationship between the percentage of items read carefully and the degree of missed instructed response items. It could be that the item about reading items carefully is more specific than the others and therefore more accurately answered (and more likely to relate to my similarly specific dependent measure of missed items), whereas the items about paying attention and percent of time spent on other tasks are more general or abstract. Regardless, study condition still affected all of the supplemental dependent measures involving self-reported attentiveness and multitasking, so all of the measures seemed to have some value in gauging the level of inattentiveness.

The increased likelihood for inattention in the non-lab condition observed in the present study highlights why a main purpose of the study was to implement a replicable procedure (instructed response items) to detect cases of inattention that may have gone unnoticed with other methods, such as outlier or response rate analyses. They also helped to establish the prevalence of inattention in my study, which was generally low, even in the non-lab online condition. Some participants had high percentages of missed instructed response items, but the vast majority of participants missed 0% of instructed response items. Instructed response items can also inform researchers of which participants' data should be removed from analyses. The inclusion of instructed response items does not seem to have any notable drawbacks that outweigh these potential advantages. Based on participant feedback, the instructed response items did not make the study any more burdensome. Only a couple of participants described the items as confusing, and the confusion pertained more to a general curiosity as to why they were included in the study (e.g., "I wondered why they were there"). For these reasons, it is concluded that instructed response items are an essential tool to include in study design. They are easily implemented and interpreted, capture problematic responding behaviors, and are non-intrusive to participants.

A primary implication from the results, despite the equivalence of paper and online formats in the lab, is that we are not yet ready to transition solely to online data collection for studies which require sustained attention, namely in uncontrolled settings. Future research should entail identifying and controlling for the barriers inherently associated with research in natural settings when it comes to online studies. For example,

one solution may to be to have open laboratories in which participants drop in and complete online studies.

Another potential direction for future research is to include cell phones as a modality that participants can use to complete the study. Cell phones were prohibited as a method of completing the current study. Participants were instructed to complete the study on paper or an electronic device other than a cell phone (e.g., desktop computer, laptop, iPad, etc.). Because this study specifically focused on establishing a novel comparison of inattention across paper and online conditions using instructed response items, it was thought that including cell phones would add an unnecessary potential confound limiting the conclusions from this initial comparison and should instead be a second step for future research. Cell phones screens vastly differ in interface and navigation from paper and computers. The current study took care to make the appearance of measures across conditions look as identical as possible to control for extraneous variables. However, since no significant differences were found between the lab paper and lab computer condition in the present study, future research may want to build upon these results by manipulating study conditions to include cell phones. This may also add to the generalizability of findings as most college students and people in general own cell phones. One participant in the current study was excluded from analyses for using a cell phone to complete the study. This participant missed a great deal of instructed items, indicating that future research may find more problematic responses on cell phones than on other electronic formats or paper.

Indeed, the lack of allowance of cell phones may be a reason for the lower than expected overall rate of inattentive responding in the present study. Another potential

reason is participant self-selection. Specifically, my study was categorized as a lab study and it was only after participants had already signed up for a lab timeslot that a portion of participants were sent the link to participate online. It is possible that people who are willing to come into the lab for a study may pay more attention to survey items overall.

The low level of inattentive responding could be considered a limitation of the study, given that our dependent measure had low variability. Another limitation of the current study was that the non-lab online condition had fewer participants than the two lab conditions. Though the current study still satisfied statistical assumptions for valid analyses, future studies may benefit from having more equal group sizes, and/or more participants in general across conditions to increase statistical power, and variety of analyses able to be conducted. Weigold et al. (2013) suggest unequal sample sizes is a common issue that may have contributed to the historically inconsistent findings of equivalence research in paper and online modalities.

## Conclusion

These findings are promising for university researchers (or otherwise) who conduct studies which require sustained attention because they reinforce that online data collection is as viable an option as paper data collection, though there are some stipulations. Based on my findings, I recommend researchers who use online data collection implement a system such as instructed response items to capture inattentive participant behavior. Researchers might also consider assessing individual difference factors that could contribute to inattentiveness. Of most importance, researchers using or interpreting online data collection should use caution when allowing participants to complete online studies in uncontrolled environments. It is hoped that further research

will identify methods for counteracting the increased inattention in online studies that take place in uncontrolled environments, so that researchers can use that convenient method of data collection confidently.

References

Alexander, P. A., & Singer, L. M. (2017, October 15). A new study shows that students learn way more effectively from print textbooks than screens. *Business Insider*. Retrieved from http://www.businessinsider.com/students-learning-education-print-textbooks-screens-study-2017-10?IR=T

Barthakur, R. (2013). Computer vision syndrome. *Internet Journal of Medical Update, 8*(2), 1–2. Retrieved from https://www.akspublication.com/Editorial_Jul2013_.pdf

Beach, D. A. (1989). Identifying the random responder. *Journal of Psychology: Interdisciplinary and Applied, 123*(1), 101–103. Retrieved from http://psycnet.apa.org/record/1989-28058-001

Berry, D. T. R., Wetter, M. W., Baer, R. A., Larsen, L., Clark, C., & Monroe, K. (1992). MMPI-2 random responding indices: Validation using a self-report methodology. *Psychological Assessment, 4,* 340–345. doi:10.1037/1040-3590.4.3.340

Birkerts, S. (1996). Reading in the electronic era. *Logos, 7*(3), 211–214.

Brown, K. W., & Ryan, R., M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84,* 822-848. Retrieved from https://selfdeterminationtheory.org/SDT/documents/2003_BrownRyan.pdf

Butcher, J. N. (1989). Minnesota multiphasic personality inventory. In *Corsini Encyclopedia of Psychology*. (pp. 1-3). John Wiley & Sons, Inc.

Carriere, J. A., Cheyne, J. A., & Smilek, D. (2008). Everyday attention lapses and memory failures: The affective consequences of mindlessness. *Consciousness and*

*Cognition: An International Journal*, *17*, 835-847. doi:10.1016/j.concog.2007.04.008

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*, 2156-2160. doi:10.1016/j.chb.2013.05.009

Chatfield, T. (2015, February). Why reading and writing on paper can be better from your brain. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2015/feb/23/reading-writing-on-paper-better-for-brain-concentration

Cheyne, J. A., Carriere, J. A., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and Cognition: An International Journal*, *15*, 578-592. doi:10.1016/j.concog.2005.11.009

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349-354. doi:10.1037/h0047358

Dias, J. C. R., Maroco, J., & Campos, J. A. D. B. (2015). Weight Concerns Scale applied to college students: Comparison between pencil-and-paper and online formats. *Cyberpsychology, Behavior and Social Networking*, *18*(3), 188–192. doi:10.1089/cyber.2014.0392

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, *49*(1), 71-75. doi:10.1207/s15327752jpa4901_13

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*(2), 192-203. doi:10.1037/1040-3590.18.2.192

Duffy, B., Smith, K., Terhanian, G., & Bremer, J. (2015). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, *47*, 615-639. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.376.621&rep=rep1&type=pdf

Eden, S., & Eshet-Alkalai, Y. (2012). Print versus digital: The effect of format on performance in editing text. In: Eshet-Alkalai, Y., Caspi, A., & Eden, S. et al. (Eds.), *Proceedings of the Chais Conference on instructional technologies research 2012: Learning in the technological era* (pp. 13-21). Raanana: The Open University of Israel.

Farmer, R., & Sundberg, N. D. (1986). Boredom proneness: The development and correlates of a new scale. *Journal of Personality Assessment, 50*(1) 4-17. doi:10.1207/s15327752jpa5001_2

Gentile, B., Miller, J. D., Hoffman, B. J., Reidy, D. E., Zeichner, A., & Campbell, W. K. (2013). A test of two brief measures of grandiose narcissism: The Narcissistic Personality Inventory-13 and the Narcissistic Personality Inventory-16. *Psychological Assessment, 25*, 1120-1136. doi:10.1037/a0033192

Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. Mervielde, I., Deary, I. J., De Fruyt, F., & Ostendorf, F. (Eds.), *Personality psychology in Europe, 7*(pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.

Gosling, S., Vazire, S., Srivastava, S., & John, O. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59*(2) 93–104. doi:10.1037/0003-066X.59.2.93.

Greszki, R., Meyer, M., & Schoen, H. (2015). Exploring the effects of removing "too fast" responses and respondents from web surveys. *Public Opinion Quarterly, 79*, 471-503. doi:10.1093/poq/nfu058

Harrington, N. (2005), The Frustration Discomfort Scale: Development and psychometric properties. *Clinical Psychology and Psychotherapy, 12*, 374–387. doi:10.1002/cpp.465

International Personality Item Pool. (2017, February 22). Converting IPIP item responses to scale scores. Retrieved from http://ipip.ori.org/newscoringinstructions.htm

Karlberg, L., Krakau, I., Sjödén, P., & Undén, A. (1997). Psychometric properties of a brief self-report Type A questionnaire for use in primary health care. *Scandinavian Journal of Primary Health Care*, *15*(1), 52-56. doi:10.3109/02813439709043431

Kopp, J. P., Zinn, T. E., Finney, S. J., & Jurich, D. P. (2011). The development and evaluation of the academic entitlement questionnaire. *Measurement and*

*Evaluation in Counseling and Development*, *44*(2), 105-129. doi:10.1177/0748175611400292

Kretzschmar, F., Pleimling, D., Hosemann, J., Füssel, S., Bornkessel-Schlesewsky, I. & Schlesewsky, M., (2013). Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media, *PLoS ONE, 8*(2), e56178. doi:10.1371/journal.pone.0056178.

Leite, W. L., & Beretvas, S. N. (2005). Validation of scores on the Marlowe-Crowne Social Desirability scale and the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement*, *65*(1), 140-154. doi:10.1177/0013164404267285

LePera, N. (2011). Relationships between boredom proneness, mindfulness, anxiety, depression, and substance use. *The New School Psychology Bulletin*, *8*(2), 15-25. Retrieved from http://www.azmonyar.com/DownloadPDF/15562289.pdf

Liu, Z. (2012). Digital reading. *Chinese Journal of Library and Information Science (English edition), 5*(1) 85-94. Retrieved from https://pdfs.semanticscholar.org/4746/aeeb8b966f8ba4520715ba8c3856bacc9c39.pdf

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, *46*(2), 137-155. doi:10.1023/A:1006824100041

Mangen, A., Walgermo, B. R., Bronnick, K. (2013) Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research, 58,* (2013) 61-68. doi:10.1016/j.ijer.2012.12.002

McLeish, K. N., & Oxoby, R. J. (2007). Measuring impatience: Elicited discount rates and the Barratt impulsiveness scale. *Personality and Individual Differences, 43,* 553-565. doi:10.1016/j.paid.2007.01.002

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17,* 437-455. doi:10.1037/a0028085

Moisala, M., Salmela, V., Hietajärvi, L., Salo, E., Carlson, S., Salonen, O., & Alho, K. (2016). Media multitasking is associated with distractibility and increased prefrontal activity in adolescents and young adults. *Neuroimage, 134,* (2016) 113-121. doi:10.1016/j.neuroimage.2016.04.011

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard. *Psychological Science, 25,* 1159-1168. doi:10.1177/0956797614524581

Myrberg, C. & Wiberg, N. (2015). Screen vs. paper: What is the difference for reading and learning? *Insights. 28*(2), 49–54. doi:10.1629/uksg.236

Nichols, D. S., Greene, R. L., & Schmolck, P. (1989). Criteria for assessing inconsistent patterns of item endorsement on the MMPI: Rationale, development, and empirical trials. *Journal of Clinical Psychology, 45*(2), 239 –250. doi:10.1002/1097-4679(198903)45:2<239::AID-JCLP2270450210>3.0.CO;2-1

Noyes, J. M., & Garland, K. J. (2003). VDT versus paper-based text: Reply to Mayes, Sims and Koonce. *International Journal of Industrial Ergonomics, 31*, 411–423. doi:10.1016/S0169-8141(03)00027-1

Oppenheimer, D., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45,* 867–872. doi:10.1016/j.jesp.2009.03.009.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51,* 768 –774. doi:10.1002/10974679(199511)51:6_768::AIDJCLP2270510607_3.0.CO;2-1

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609. doi:10.1037/0022-3514.46.3.598

Raskin, R. & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports, 45,* 590. doi:10.2466/pr0.1979.45.2.590

Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54,* 890-902. doi:10.1037/0022-3514.54.5.890

Sanchez, C. A., & Wiley, J. (2009). To scroll or not to scroll: Scrolling, working memory capacity, and comprehending complex texts. *Human Factors, 51,* 730-738. doi:10.1177/0018720809352788.

Scripture, E. (1893). Psychological Measurements. *The Philosophical Review, 2,* 677-689. doi:10.2307/2176022

Sherman, E. D., Miller, J. D., Few, L. R., Campbell, W. K., Widiger, T. A., Crego, C., & Lynam, D. R. (2015). Development of a Short Form of the Five-Factor Narcissism Inventory: The FFNI-SF. *Psychological Assessment, 27,* 1110-1116. doi:10.1037/pas0000100

Smith, M. M., Saklofske, D. H., Stoeber, J., & Sherry, S. B. (2016). The Big Three Perfectionism Scale: A new measure of perfectionism. *Journal of Psychoeducational Assessment, 34,* 670-687. doi:10.1177/0734282916651539

Steinberg, L., Sharp, C., Stanford, M. S., & Tharp, A. T. (2013). New tricks for an old measure: The development of the Barratt Impulsiveness Scale–Brief (BIS-Brief). *Psychological Assessment, 25*(1), 216-226. doi:10.1037/a0030550

Thomas, M., & Watkins, P. (2003, May). *Measuring the grateful trait: Development of the revised GRAT*. Presentation to the 83rd annual convention of the Western Psychological Association, Vancouver, British Columbia.

Tourangeau, R. (1984). Cognitive science and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), Cognitive aspects of survey methodology: Building a bridge between disciplines (pp. 73- 100). Washington, DC: National Academy Press.

U.S. Census Bureau. (2014). American community survey. Retrieved from https://www.census.gov/library/visualizations/2016/comm/digital_nation.html

U.S. Census Bureau. (2014) Computer and internet use in the United States: 2013. Retrieved from https://www.census.gov/library/publications/2014/acs/acs-28.html

Vallerand, Robert J., Pelletier, Luc G., Blais, Marc R., Brière, Nathalie M., Senecal, Caroline, & Vallieres, Evelyne F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*, 1003-1017. doi:10.1177/0013164492052004025

Watkins, P. C., Woodward, K., Stone, T., & Kolts, R. L. (2003). Gratitude and happiness: Development of a measure of gratitude and relationships with subjective well-being. *Social Behavior and Personality, 31*, 431-452. doi:10.2224/sbp.2003.31.5.431

Sanchez, C. A., & Wiley, J. (2009). To scroll or not to scroll: Scrolling, working memory capacity, and comprehending complex texts. *Human Factors, 51*, 730-738. doi:10.1177/0018720809352788.

Weigold, A., Weigold, I. K., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. *Psychological Methods, 18*(1), 53-70. doi:10.1037/a0031607

Wood, B., Rea, M. S., Plitnick, B., & Figueiro, M. G. (2013). Light level and duration of exposure determine the impact of self-luminous tablets on melatonin suppression. *Applied ergonomics, 44*(2), 237-240. doi:10.1016/j.apergo.2012.07.008

Woolliscroft, J. O., TenHaken, J., Smith, J., & Calhoun, J. G. (1993). Medical students' clinical self-assessments: comparisons with external measures of performance and the students' self-assessments of overall performance and effort. *Academic Medicine, 68*(4), 285-94. doi:10.1097/00001888-199304000-00016

## Appendix A

*List of Instructed Response Items by Measure*

| Measure | Instructed Response Item |
|---|---|
| TASK 1<br>Balanced Inventory of Desirable Responding | None |
| TASK 2<br>Frustration Discomfort Scale | (1) For data checking purposes, please choose option 3, Moderately agree, for this item in order to assist the researchers conducting this study.<br>(2) For data checking purposes, please choose option 4, Strong, for this item. |
| TASK 3<br>Gratitude Resentment and Appreciation Test | (3) For this item, please select option 8 to assist the researchers of this study.<br>(4) Choose option 2 for this item for the purpose of data checking.<br>(5) Please select option 4 to assist the researchers of this study in data checking.<br>(6) Please select option 5, Feel neutral, for data checking purposes.<br>(7) Choose option 1, Strongly disagree, to assist the researchers of this study. |
| TASK 4<br>Academic Motivations Scale-College | (8) For data checking purposes, please select option 7, Agree completely.<br>(9) For data checking purposes, please select option 7, Agree completely. |
| TASK 5<br>Simplified 'Type A' Questionnaire | (10) Choose Do not agree at all, for data checking purposes. |
| TASK 6<br>Academic Entitlement Questionnaire | (11) To assist the researchers of this study, choose option 7, Strongly agree. |
| TASK 7<br>Attention-Related Cognitive Errors Scale | None |
| TASK 8<br>Narcissistic Personality Inventory | (12) Please choose B as your answer for this item.<br>(13) For data checking purposes, please select B.<br>(14) Select A for this item for data checking purposes.<br>(15) Choose A to assist with data checking. |
| TASK 9<br>Boredom Proneness Scale | (16) For data checking purposes, select option 2 for this item. |

| | |
|---|---|
| TASK 10<br>Subjective Happiness Scale | None |
| TASK 11<br>Mindful Attention Awareness Scale | (17) Choose option 3, Somewhat frequently, for this item to assist the researchers of this study. |
| TASK 12<br>Big Three Perfectionism Scale | (18) For data checking purposes, please choose option 1, Strongly disagree.<br>(19) Please choose option 4 for data checking.<br>(20) Select option 5, Strongly agree, for this item. |
| TASK 13<br>Mini International Personality Item Pool | None |
| TASK 14<br>Satisfaction with Life Scale | None |
| TASK 15 Barratt Impulsiveness Scale | (21) Please choose option 4, Almost always/Always, for this item for data checking purposes.<br>(22) Choose Occasionally to assist the researchers of this study with data checking. |

## Appendix B

*Supplemental Questions*

| No. | Item |
|---|---|
| 01. | Roughly how many psychological studies have you participated in before this one? |
| 02. | What is your current GPA? |
| 03. | What type of device did you use to complete this study? (e.g., paper, Mac laptop) |
| 04. | What location were you in when you participated in this study? (e.g., home, EWU lab) |
| 05. | Please estimate your average digital screen time per day (e.g., 6 hours and 30 minutes). This includes all of the time you spend on desktops, laptops, iPads, cell phone, etc. |
| 06. | Did anything disrupt your reading of the survey? If yes, briefly describe. |
| 07a. | What percentage of time did you spend on other tasks while taking the survey? Please enter exact number between 0 and 100. |
| 07b. | What activity/activities were you participating in when you were multitasking (e.g., texting)? If you were not multitasking, please write "N/A". |
| 08. | The survey was designed to take approximately 45 minutes to complete, but, without checking the time, how long did it feel like it took (i.e., subjective time)? Please provide exact number of minutes. |
| 09. | What do you believe was the purpose of this study? |
| 10. | Did you find anything about this study to be unusual, confusing or suspicious? If yes, please explain. |
| 11. | Did inclusion of data-checking items (e.g., "Please select option 3, Strongly disagree, for data checking purposes) affect how you answered the survey? Explain. |

| 12. | | Strongly disagree (1) | (2) | (3) | (4) | (5) | (6) | Strongly agree (7) |
|---|---|---|---|---|---|---|---|---|
| | I prefer reading electronic textbooks over reading paper textbooks. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| 13. | | Strongly disagree (1) | (2) | (3) | (4) | (5) | (6) | Strongly agree (7) |
|---|---|---|---|---|---|---|---|---|
| | I prefer reading electronic articles rather than articles printed on paper. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

14.

| | Strongly disagree (1) | (2) | (3) | (4) | (5) | (6) | Strongly agree (7) |
|---|---|---|---|---|---|---|---|
| I pay more attention to tasks when they're on paper compared to when they're electronic. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

15.

| | Not at all boring (1) | (2) | (3) | (4) | (5) | (6) | Extremely boring (7) |
|---|---|---|---|---|---|---|---|
| How boring was this study? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

16.

| | Not at all (1) | (2) | (3) | (4) | (5) | (6) | A great deal (7) |
|---|---|---|---|---|---|---|---|
| To what extent did you pay attention while completing the study? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

17.

| | Strongly disagree (1) | (2) | (3) | (4) | (5) | (6) | Strongly agree (7) |
|---|---|---|---|---|---|---|---|
| Psychological research interests me. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

18.

| | Not at all (1) | (2) | (3) | (4) | (5) | (6) | A great deal (7) |
|---|---|---|---|---|---|---|---|
| To what extent do you value social science research? | ○ | ○ | ○ | ○ | ○ | ○ | |

19. Approximately what percent of items did you read carefully? Please give an exact percentage between 0 and 100.

20.

| | Extremely unlikely (1) | Moderately unlikely (2) | Slightly unlikely (3) | Neither likely nor unlikely (4) | Slightly likely (5) | Moderately likely (6) | Extremely likely (7) |
|---|---|---|---|---|---|---|---|
| How likely do you believe it is that this study would lead to research that gets published in a scientific journal? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

21.

| | Extremely unlikely (1) | Moderately unlikely (2) | Slightly unlikely (3) | Neither likely nor unlikely (4) | Slightly likely (5) | Moderately likely (6) | Extremely likely (7) |
|---|---|---|---|---|---|---|---|
| How likely do you believe it is that this study would lead to research that gets published in a textbook? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 22. | | Extremely unlikely (1) | Moderately unlikely (2) | Slightly unlikely (3) | Neither likely nor unlikely (4) | Slightly likely (5) | Moderately likely (6) | Extremely (7) | |
| | How likely do you believe it is that this study would lead to research that gets presented at a conference? | ○ | ○ | ○ | ○ | ○ | ○ | ○ | |
| | | | | | | | | | |
| 23. | If you took this study outside of the lab setting, did you complete the survey all at once, or did you leave and come back? Please explain. | | | | | | | | |

# VITA

Author: Olivia B. Brooks

Place of Birth: Greenville, South Carolina

Undergraduate Schools Attended: University of Montana

Degrees Awarded: Bachelor of Arts, 2013, University of Montana

Honors and Awards: Graduate Service Appointment, Eastern Washington University, 2015-2016

Graduated Magna Cum Laude, University of Montana, 2013

Psi Chi Honor Society

Alpha Lambda Delta Honor Society

Professional Experience: Occupational Sleep Medicine Technical Assistant, Washington State

University Spokane, College of Medicine, Sleep and Performance Research Center, Spokane, WA, 2017-2018

Chemical Dependency Professional Trainee (CDPT), Pioneer Human Services, Spokane, WA, 2017

Patient Services Technician, Pioneer Human Services, Spokane, WA, 2017

Behavioral Health Intern, Community Health Association of Spokane (CHAS), Spokane and Cheney, Washington, 2016-2017

Clinical Research Coordinator, Washington State University Spokane, College of Nursing, Program of Excellence in Addictions Research (PEAR), Spokane, WA, 2014-2015

Research Assistant, Washington State University Spokane, College of Nursing, Program of Excellence in Addictions Research (PEAR), Spokane, WA, 2013-2014