12-12-2019

Seduction Is in the Eye of the Beholder: The Influences of Learner Characteristics on Learning from Multiple Texts with Seductive Details

Alex August Johnson

SEDUCTION IS IN THE EYE OF THE BEHOLDER: THE INFLUENCES OF LEARNER CHARACTERISTICS ON LEARNING FROM MULTIPLE TEXTS WITH SEDUCTIVE DETAILS

by

Alexander A. Johnson

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Major: Psychology

The University of Memphis

December 2019

Acknowledgements

Science is not conducted in a vacuum and this project is no exception. Many wonderful people have assisted with this project and inspired and supported me along the way. First and foremost, I would like to thank my advisors: Dr. Roger Kreuz's guidance with writing and editing this manuscript was invaluable and Dr. Jason Braasch's assistance with developing materials, recruiting participants, and general advice were essential to this project. I would also like to thank Dr. Dorian Burnette for taking the time to serve on this committee.

Further, I am grateful for the feedback and suggestions provided by the faculty in the Cognitive Psychology research area at the University of Memphis and for the insight and knowledge I have gained from working with them. A particular thank you is owed to Dr. Andrew Olney for his quantitative advice on portions of this project and in general. I would also like to thank the research assistants from the Knowledge Acquisition Lab who kindly gave their time and whose diligence was extremely helpful for data management.

Additionally, I am grateful for the Institute for Intelligent Systems and the Institute for Intelligent Systems Student Organization's Thesis Award, which provided financial support for the pilot study for this project, as well as for the variety opportunities this organization has provided me.

Finally, I would like to thank my family, friends, and fiancé for their encouragement throughout this project and my education, generally. I am truly grateful for all the support I have received.

Abstract

Previous research has shown that the inclusion of *seductive details*—highly interesting, yet conceptually unimportant information—in educational materials is detrimental to learning. Recently, however, researchers have suggested that learner characteristics and motivational factors may influence these detriments. This study built on these findings by exploring the effects of seductive details on multiple-text comprehension, as well as including measures that have not been explored previously (e.g., need for cognition). Additionally, reading and response times were collected to examine online processes during comprehension. Participants were asked to read two short texts about El Niño, sentence-by-sentence, and were later asked to judge whether provided inferences could be drawn by connecting information between texts (i.e., the IIVT), as well as to provide their confidence in these judgments. Those receiving a text with seductive details performed worse on the IIVT. Additionally, both vocabulary knowledge and need for cognition were significant predictors of IIVT performance.

Table of Contents

List of Tables

Introduction

In 1913, John Dewey published *Interest and Effort in Education*, in which he argued for the importance of interest in education, arguing that "if we can secure interest in a given set of facts or ideas, we may be perfectly sure that the pupil will direct his energies toward mastering them" (Dewey, 1913/2012, p. 1). Over 100 years later, educators are still grappling with this task. Instructors, textbook writers and editors, and educational software companies have attempted to increase how interesting and engaging their materials are. Instructors may attempt to incorporate interesting activities or videos, textbook editors may insert entertaining anecdotes about a topic, and companies increasingly attempt to create more aesthetically pleasing multimedia resources for learning. Given this emphasis on increasing the interestingness of educational materials, it is worth asking the question: are we *actually* improving learning?

Without a doubt, articles that are more vivid can increase interest in texts (e.g., Wade, Buxton, & Kelly, 1999), multimedia learning can lead to better learning outcomes (e.g., Mayer, 1997), and classroom activities may get students more engaged (e.g., Cavanagh, 2011). However, through attempts to increase the interestingness of texts, a paradoxical detrimental effect of interesting materials has been revealed, referred to as the seductive details effect (SDE). A large body of literature has shown that while increased interest can lead to improved learning outcomes, the way that interest is increased matters greatly. The SDE shows that increasing interest through the addition of highly interesting, *but irrelevant*, information actually has a detrimental effect on learning (e.g., Garner, Gillingham, & White, 1989; Mayer, Heiser, & Lonn, 2001; Rey, 2012; Sanchez & Wiley, 2006).

Results have overwhelmingly supported the idea that the inclusion of these highly interesting but irrelevant details has negative effects on learning (e.g., Rey, 2012). Recently,

however, boundary conditions on this effect have been observed and several moderators have been found to influence this effect (e.g., Korbach, Brünken, & Park, 2016). While the detrimental effects of seductive details still hold true, it seems that they are not uniformly harmful, and that some may even benefit from these distracting pieces of information (e.g., Wang & Adesope, 2016). While researchers are not advocating for the inclusion of these details, it is clear that the SDE may be more nuanced. Based on these recent findings, the current project sought to explore how certain learner characteristics may protect against the SDE. Further, I hoped to test competing hypotheses regarding what causes this effect.

Text Comprehension

The SDE is one finding seated within the broader text comprehension literature. Comprehension refers to the processing of information to extract meaning and is necessary for nearly all higher-level processes, such as learning, reasoning, or decision making (McNamara & Magliano, 2009). While this area of research is quite vast and complex, several key ideas will be discussed that are of relevance for the current project, including a general text comprehension model, how these processes can be affected by characteristics of a text (e.g., coherence), and how individuals comprehend multiple texts.

There are many models of text comprehension, which may vary depending on the outcome of interest for the researcher (e.g., recall, problem-solving), the types of text being discussed (e.g., narrative, expository), or other points of emphasis. McNamara and Magliano (2009) reviewed dominant theories of comprehension and found several areas of overlap between them. For example, models shared a view that comprehension involved a connectionist architecture where words, the current context, underlying meaning, and prior knowledge are activated concurrently. Spreading activation leads to activation of concepts related to those being

read about. Further, when cohesion gaps are encountered, inferences may be generated which then become part of the model of the text. There are other similarities as well, but these are largely unimportant for the current work. Instead, one model will be described, as it provides sufficient information for this project and differences from other models involve nuanced distinctions that would not affect interpretation of the SDE.

The Construction-Integration Model

The Construction-Integration (CI) model was proposed by Kintsch in 1988 and still is in general alignment with more contemporary theories, which are often clarifications or extensions of this model. *Construction* in this model refers to the initial activation of the information in the text with existing related knowledge. Activation can come from the current input (e.g., sentence), previous information from the text, and general related knowledge. *Integration* refers to the eventual settling of these memories as spreading activation decreases. That is, information is activated and new ideas can be incorporated into existing networks or relationships between existing networks are strengthened. Information that is repeatedly activated is strengthened in memory (e.g., *genetics* in a paper on evolution), while tangentially related information loses its strength as the lack of repeated activation of these tangential concepts leads to weakening connections (e.g., *whale*, mentioned only once, in a paper on evolution). Through processes beyond the scope of this paper (see McNamara & Magliano, 2009, for review), these networks of activation are stabilized.

The CI model posits that sentences lead to three levels of representation: a *surface structure* representing the words in a text and syntax, a *propositional textbase* consisting of sets of propositions, and a *situation model* which refers to inferences and information beyond what is explicitly mentioned (McNamara & Magliano, 2009). Different forms of testing are suggested to

reflect these other components of comprehension. For example, recall (e.g., short-answer questions) is likely to reflect the textbase model, while problem-solving questions or essay questions are more likely to reflect the situational model. For our purposes, the situation model is most relevant, as it represents the model of the text formed through its connections with prior knowledge. This deeper, more integrated representation of the concept is what educators generally hope to construct, as well as what many researchers use to examine comprehension. To foreshadow, this study used a measure intended to represent learners' ability to form inferences between two texts, reflecting their situation model level of comprehension.

Multiple-Text Comprehension

When individuals read multiple texts on a topic, more factors must be considered in the formation of a coherent representation of the information. That is, simple representations of sequential texts are rarely adequate. For example, Britt, Rouet, and Durik (2018) suggest that readers need to represent multiple texts distinctly while also representing their interrelated components, such as areas of agreement or disagreement. Several models have been proposed to explain the processes underlying multiple-text comprehension. For the purposes of this paper, the Documents Model Framework (DMF) provides sufficient information as the current study involves two complementary texts with similar form and no source information (e.g., author) available. That is, while the strengths of more contemporary models (e.g., RESOLV; Britt et al., 2018) are acknowledged, the additional motivational considerations are beyond the scope of this project and represent a difference in focus and clarification of additional influences, rather than a contradiction of the components represented by DMF (e.g., Bråten, Braasch, & Salmerón, in press).

According to the DMF, readers will ideally create two structures beyond those included in single text comprehension. First, they should create an *integrated mental model* which represents a global model of information described in the texts. This may include information from single texts, repeated ideas, or contradictory information. Additionally, they should create an *intertext model* which uses source information (e.g., authors) to organize their representation. Ideally, this allows connections between texts for relationships, such as the source of contrasting or supplemental views on a topic (e.g., Bråten, Braasch, & Salmerón, in press). In the current study, two texts are used with no source information available. As such, it is likely that representations may lack some of these intertext connections. However, participants will also not be tested on their memory for the sources. Instead, our final test will measure their ability to form inferences across texts, reflecting a test that better targets readers' integrated mental models.

Inferences from Text

Beyond remembering information contained within a text, readers also engage in various forms of inference construction during reading. For example, readers may use inferences to fill in gaps in the text by connecting information with previous portions of the text—sometimes referred to as *local cohesion inferences* or *bridging inferences*—or by connecting it with previous knowledge on the topic or related information—sometimes referred to as *global coherence inferences* or *elaborations* (e.g., Cain & Oakhill, 2014). Exactly how inferences are generated and the conditions under which certain types are generated is subject to many areas of disagreement (e.g., Cook & O'Brien, 2016). As such, this study takes a theoretically naïve position and asserts only that inferences are made, that they can involve prior knowledge, and that they are useful for more intricate and thorough understandings of a text.

In many areas of education, inferences are extremely useful for a more thorough understanding of to-be-learned information. For example, connecting the information about "hedonic hotspots" in the nucleus accumbens (e.g., Peciña, Smith, & Berridge, 2006) to clinical studies of addiction may allow individuals to better understand how the brain plays a role in this process. In the current study, the ultimate goal was for participants to learn how different elements of weather systems (e.g., air pressure, temperature) contribute to weather patterns across the Pacific Ocean during typical and El Niño Southern Oscillation cycles. In order to form a more complete understanding of the text, participants had to form inferences between the two texts. The primary measure of learning was a test of participants' ability to draw inferences from the two texts they read. Specifically, they were asked to judge whether a provided inference could be made by combining information contained across two texts, a measure referred to as the Intertextual Inference Verification Task (IIVT).

Previous research on inferences has shown that several of the measures included in the current study should be related to their performance on the IIVT. For example, a study on adolescents by Cain and Oakhill (2014) showed that vocabulary knowledge was a stronger predictor of these deeper, global inferences, than of local inferences or literal memory for the text. As Cain and Oakhill tested these ideas on adolescents, their relevance for the current study is reduced. However, it does lend conceptual support for the particular relevance of vocabulary in making inferences and is consistent with other studies of vocabulary's contribution to inferencing in children (e.g., Currie & Cain, 2015). Additionally, research using adult samples like those used in the current study has shown consistent patterns (e.g., Calvo, 2004; Singer, Andrusiak, Reisdorf, & Black, 1992).

In the current study, the inferences tested by the IIVT likely fall in between these two extremes—certainly, it is necessary to combine prepositions from across texts to determine the validity of the tested inferences; however, determining whether or not these inferences are valid is also influenced by prior knowledge and, as such, may reflect more global inferences. Regardless of which form of inference is more relevant, however, consistent findings supporting the role of vocabulary knowledge on inference ability lead to the prediction that vocabulary knowledge will be associated with improved IIVT performance in the current study.

H_1: Vocabulary knowledge will be associated with improved IIVT performance.

In an eye-tracking study by Calvo (2004), vocabulary knowledge was associated with reduced regressions (i.e., re-reading portions of a text, often associated with attempts to ensure comprehension) in an inference task, suggesting that individuals with higher vocabulary knowledge exerted less effort in forming these inferences. It is worth noting that generally high levels of performance may have prevented any effects of vocabulary knowledge on comprehension and inference accuracy to be inferred. However, support from online measures (i.e., eye-tracking) further supports the role of vocabulary knowledge in inferencing. Additionally, albeit indirectly, this suggests that improved vocabulary knowledge may decrease the load imposed on limited cognitive resources during processing, thus increasing the probability of an inference being made (e.g., Singer et al., 1992). These studies support my prediction that vocabulary knowledge will be a significant predictor of IIVT performance.

Prior Knowledge

While some implications of the effects of prior knowledge on learning are obvious—for example, that individuals who previously learned something will recall it better when they

encounter it a second time—there are some less intuitive, but theoretically important, contributions of prior knowledge. As previously described, prior knowledge can facilitate reading comprehension as the reader has a greater amount of background information that increases how easily this new information is retrieved. For example, a physicist trained in aerodynamics could connect information about pressure systems in weather to his or her existing knowledge, improving memory for weather systems by linking it to an existing strong and conceptually related representation. Similarly, beyond merely connecting this information to a stronger memory, an individual with higher prior knowledge typically has a more coherent understanding.

In addition to these more memorial effects, prior knowledge can influence the processing of the text. For example, McNamara et al. (1996) found that prior knowledge interacts with text cohesion in comprehension. Text *cohesion* refers to the degree to which ideas presented in a text are made explicit and clear by including features that guide the reader in understanding ideas, connecting ideas within the text, and connecting textual ideas to larger themes (Graesser, McNamara, Louwerse, & Cai, 2004). Graesser et al. (2004) distinguish this from *coherence*, which they explain is a characteristic of the reader's mental representation of the text. While these are distinct phenomena, then, they are interrelated in that text cohesion may increase the coherence of the reader's understanding.

In a study by McNamara et al. (1996) middle and high school students with high and low prior knowledge read texts that differed in coherence. They were then given several tests, including recall questions based on information from single sentences and inference questions that required combining information from multiple sentences in the text. Individuals with high prior knowledge performed better on the inference questions when they read a low coherence

text. Conversely, those with low prior knowledge benefitted from a more coherent text. They explained this pattern by arguing that low coherence texts required readers to fill in gaps in the information, which high prior knowledge students were able to do as they possessed the information necessary to fill these gaps. Conversely, low prior knowledge students did not have the information necessary, so their understanding and corresponding ability to solve inference problems was worse (McNamara et al., 1996). This is only one study but illustrates that prior knowledge can influence text comprehension in complicated ways. To foreshadow, it also introduces one factor that may explain why seductive details are not always detrimental (i.e., coherence disruption may benefit some learners).

Based on previous research examining the influence of prior knowledge on inference generation—such as the findings described from McNamara et al. (1996)—it is expected that individuals with high prior knowledge will perform better on our inference-based measure of learning, the IIVT.

H_2: Participants with high prior knowledge will perform better on the IIVT than those with low prior knowledge.

Interest, Broadly Speaking

As mentioned previously, one goal among educators, editors, and authors is to create learning materials that create and maintain student interest. However, this begs the question: What is interest? Psychologists have discussed the role of interest since the emergence of the field, including references by James, Piaget, and other pioneering figures; however, it has only been studied in a more structured and systematic sense in the past several decades. As seductive

details are intended to increase the interestingness of texts, the idea of interest itself warrants a more thorough discussion.

Although there are many differences in how interest is conceptualized and measured, Renninger and Hidi (2011) pointed out five characteristics that most researchers agree on. First, interest is content or object specific and refers to individuals' focused attention and interactions with these events or objects. Second, it is a relationship sustained by the interaction between individuals and their environment. Individuals possess a potential intrinsic predisposition toward certain interests, such as higher spatial reasoning ability, and the content and context associated with experiences of these potential interests determines their development. Third, it includes both cognitive and affective components, though this likely varies as this interest is developed. Fourth, individuals may not even be aware of their interest being triggered, and those with a more well-developed interest may be so absorbed in their task that they may not be aware of their interest while performing it. Finally, interest has a neurological basis with distinct patterns of activation seen when a learner is or is not interested. Specifically, it is argued to be related to reward processes (Renninger & Hidi, 2011).

Interest is not a unified construct, however, and there have been many distinct approaches to viewing and studying it. For instance, some have examined interest primarily as an emotion (e.g., Silvia, 2008), while others have emphasized the features of the text or task instructions (e.g., Schraw & Lehman, 2001). While interest is likely a combination of all these components, the current review will focus on Hidi and Renninger's *Four-Phase Model of Interest Development*, as it incorporates many of these ideas into a broader framework viewing interest as a series of processes representing a continuum of interest development (Hidi & Renninger, 2006). As such, the discussion will focus on those ideas most relevant for this model, though

areas of importance that have received less attention through this model will also be reviewed, such as the knowledge deficit explanations provided by Rotgans and Schmidt (2014).

In most models of interest, there is a distinction between *individual interest*—also referred to as *personal interest*—and *situational interest*—sometimes referred to simply as *interest*, unqualified—representing two fundamentally different, albeit related phenomena (e.g., Ainley, 2017; Hidi & Renninger, 2006; Krapp, 2007; Renninger & Hidi, 2011; Rotgans & Schmidt, 2017; Schiefele, 1996; Schraw, 1997). Some have argued for further sub-divisions of these concepts, such as *topic interest* (e.g., Ainley, Hidi, & Berndoff, 2002). As this paper does not intend to exhaustively cover all theories of interest, it is worth noting this similarity.

Situational Interest

Situational interest is often described as an in-the-moment state, generally triggered by external features, and which lasts for short periods of time (e.g., Ainley, 2007; Krapp, 2007; Schraw & Lehman, 2001). It is also often considered to be related to focused attention and engagement and involves some affective response (e.g., Ainley et al., 2002; Hidi & Renninger, 2006). This form of interest aligns greatly with the emotional theories, which view interest as more of a response to stimuli or a situation than to a characteristic of the individual (e.g., Silvia, 2008). A critical component of situational interest is that it is generally elicited by some aspect of the environment. Schraw and Lehman (2001) classified these interest-triggering features as *text-based*, *task-based*, or *knowledge-based*. For the purposes of this project, only text-based and knowledge-based features are relevant.

Text-based features. Text-based interest refers to features of the text itself which contribute to triggering situational interest. Hidi (1990) proposed that this includes two classes of

features, including more structural components such as how novel or surprising ideas are, as well as content-based features, including universally interesting ideas such as life themes, romance, or violence. Shraw and Lehman (2001) combined these concepts with other suggestions and delineated text-features into three subcategories: *seductiveness*, *vividness,* and *coherence*.

Seductiveness. Seductiveness refers to highly interesting but unimportant text segments (i.e., seductive details) that distract readers from more important features. Given its extensive coverage elsewhere in this paper, it is not elaborated upon here.

Vividness. Vividness refers to text segments which create suspense, surprise, or other features that capture readers' attention (Shraw & Lehman, 2001). Previous research supported a significant relationship between vividness and different measures of interest (i.e., *perceived interest*), with a correlation of $r = .40$ (Shraw, Bruning, & Svoboda, 1995). In reality, this estimate is somewhat conservative, as their measure of vividness was highly correlated ($r = .41$) with *engagement*—"the degree to which the text was thought provoking, stimulating, and timely" (Shraw et al., 1995, p. 3)—which would be subsumed under the broader definition of vividness described by Shraw and Lehman (2001). Unlike seductive details, no evidence suggests that vividness leads to detrimental effects in learning.

Coherence. Coherence describes factors that influence how easily a reader can organize main ideas in a text (Shraw & Lehman, 2001). This is also related to Shraw et al.'s (1995) dimensions of *cohesion*—"… the text's organization and clarity" (p. 3)—and *ease of comprehension*—"…whether the text was easy to remember and concentrate on" (p. 3)—which were highly correlated ($r = .63$) and associated with perceived interest as well ($r = .45$ and $r = .61$, respectively). Wade, Buxton, and Kelly (1999) also found that poor coherence led to decreased interest. Like vividness, strong coherence is not associated with any negative effects

on learning and comprehension, representing another alternative to seductive details for increasing interest in a text, without the potential for negative consequences.

Knowledge-based features. Knowledge-based features refer to the effects of prior knowledge on situational interest (Shraw & Lehman, 2001). It has traditionally been shown that this knowledge is more strongly related to individual interest, though there is some support for its influence on situational interest. Kintsch (1980) suggested that prior knowledge is related to situational interest in an inverted U-shaped function, where very low or high levels of knowledge lead to low interest. However, moderate levels of knowledge are proposed to lead to greater interest, as it creates a desire to learn more about the topic.

Individual Interest

For the purposes of this paper, the term *individual interest* is used, in alignment with Hidi and Renninger's (2006) model, as it is the framework which guided the materials used for the current study. Individual interest refers to "… a person's relatively enduring predisposition to reengage particular content over time as well as to the immediate psychological state when this predisposition has been activated" (Hidi & Renninger, 2006, p. 113). It is worth noting that despite differing terminology, this definition closely aligns with Schraw and Lehman's *personal interest*, "… information that is of enduring personal value, activated internally, and topic-specific" (2001, p. 28), and is related to Ainley et al.'s *topic interest*, an "… interest elicited by a word or paragraph that presents the reader with a topic" (2002, p. 546).

Silvia's (2008) conceptualization may also fit here, as he views interest as an emotion involving two appraisals: an evaluation of the event's novelty, and an evaluation of its comprehensibility. To find something interesting, it must be stimulating or new, and the learner

must have some foundation upon which to assess it (e.g., interest in a painting after an art class). While this emotional view does provide a useful approach to understanding interest, it is not as directly connected to learning.

In general, individual interest is more aligned with interest in a conventional sense (e.g., being *interested* in geology), and leads to different effects on learning. Individual interest facilitates intrinsic motivation and can assist in attention, self-regulation, and achievement (e.g., O'Keefe, Horberg, & Plante, 2017). For example, while triggering situational interest in a class through the description of an interesting experiment may lead to students paying attention, individual interest does not require *interesting materials* to motivate. Instead, it facilitates the intrinsic motivation that allows individuals to push through, or even enjoy, reading dry and verbose theoretical works on the topic of interest. While individual interest is highly related to, and often elicits, motivation and engagement, it is important to note that it is a distinct construct (O'Keefe et al., 2017).

Topic Interest

The idea of topic interest has been used by many researchers as well, though it lies in a somewhat distinct space from situational or individual interest, and may involve both (e.g., Renninger & Hidi, 2016). Topic interest can refer to interest elicited through text or an intrinsic predisposition (e.g., being *interested* in something). It has been used in many studies, as well (e.g., Baldwin, Peleg-Bruckner, & McClintock, 1985; Boscolo & Mason, 2003; Clinton & van den Broek, 2012; Schiefele, 1996). In fact, this form of interest may be captured by most common measures of interest, such as questions asking about students' preferences for certain topics (e.g., mental health versus visual perception).

Unfortunately, though definitions do exist for this form of interest, these are more descriptive and do not neatly align with the research by those who study interest development. For example, Ainley and colleagues define topic interest as "the interest elicited by a word or paragraph that presents the reader with a topic" (Ainley et al., 2002, p. 546). In this conceptualization, it seems more like situational interest, the fleeting experience of interest driven by external features. Ainley et al. (2002) further argue that topic interest is a more combinatorial interest, reflecting individual interest in the topic or, if an existing interest is not present, it may be more influenced by situational factors. However, while commonly used, topic interest does not seem to reflect distinctive processes, nor does it lend itself to generating novel predictions or clarifying interpretations.

Despite questions about consistency in the use of the term, topic interest has been proposed to be closely linked with inference generation. Clinton and van den Broek (2012) conducted a study to determine which cognitive mechanism underlies the positive association between topic interest and learning. They suggested that topic interest leads individuals to generate more inferences due to differences in their desired level of comprehension for the information they are reading, referred to as *standards of coherence*. They sought to demonstrate that those with higher topic interest would desire deeper comprehension of the text and generate additional inferences to achieve this goal. In support of this, they found that the relationship between topic interest and final test performance was partially explained by inference generation. Despite the use of topic interest (i.e., as opposed to individual interest), these findings support the prediction that interest should be associated with improved inference ability.

H_3: There will be a main effect of interest, with participants with higher individual interest performing better on the IIVT, in general.

The Four-Phase Model of Interest Development

Despite the breadth of research on the concept of interest, thorough theoretical frameworks are much less common and tend to vary with how interest is viewed. Again, for the current project, Hidi and Renninger's Four-Phase Model of Interest Development (2006) is emphasized as this framework has led to the development of measures of interest development (Wang & Adesope, 2016), combines findings from many areas of research on interest, and emphasizes interest in the context of learning.

According to this model, both situational and individual interest consist of two phases. In situational interest, interest is first triggered by something (e.g., an interesting idea), but can also be maintained. In individual interest, one begins with an emerging individual interest, which can then develop into well-developed individual interest. The phases of situational interest can lead to the development of an emerging individual interest, which can then develop further. Early in interest development, interest is associated with focused attention and positive affect. More developed interest is associated with these feelings of enjoyment, but also includes stored knowledge and a sense of value in the interest (Hidi & Renninger, 2006).

Phase 1: Triggered situational interest. *Triggered situational interest* refers to "a psychological state of interest that results from short-term changes in affective and cognitive processing" (Hidi & Renninger, 2006, p. 114). It is externally sparked, often by novel or exciting stimuli, or more interesting task instructions—as in Schraw and Lehman's task-based features (2001). This phase likely aligns with the more emotion-related conceptualizations of interest (e.g., Silvia, 2008). It is also a precursor to subsequent stages of interest although, importantly, it may still be elicited in subsequent phases of interest (e.g., well-developed individual interest).

Phase 2: Maintained situational interest. *Maintained situational interest* refers to "a psychological state of interest that is subsequent to a triggered state, involves focused attention and persistence over an extended episode in time, and/or reoccurs and again persists" (Hidi & Renninger, 2006, p. 114). Maintained interest, then, is a continuation of triggered situational interest beyond an initial fleeting state, but is still largely driven by external conditions, such as Schraw and Lehman's text-based or task-based features (2001). This may then lead to the development of more stable individual interest.

Phase 3: Emerging individual interest. *Emerging individual interest* is the first stage of individual interest and refers to "a psychological state of interest as well as to the beginning phases of a relatively enduring predisposition to seek repeated reengagement with particular classes of content over time" (Hidi & Renninger, 2006, p. 114). In this phase of interest development, the individual generally has some prior knowledge or previous experience which motivates them to reengage with certain topics. It is also here the student begins to value opportunities to engage with this topic, leading to a preference for choosing this topic. Importantly, in this phase the motivation begins to be internally generated, with students thinking about the topic and generating their own questions and curiosity, allowing them to persist even in more demanding tasks (Hidi & Renninger, 2006).

Phase 4: Well-developed individual interest. *Well-developed individual interest* is the final phase of interest development, representing the same psychological state as emerging individual interest, but in an enduring form. The student strongly values the opportunity to engage with this topic even more than in previous phases, tends to be more resourceful in approaching more difficult tasks, and produces "effort that feels effortless" (Hidi & Renninger,

2006, p. 115). It also allows more sustained long-term creative efforts, facilitates deeper learning, and allows perseverance even in the face of frustrating tasks.

Wang and Adesope (2016) developed a series of measures reflecting the progression through these four phases of interest development and applied them to a study of the seductive details effect. These measures were used for the current study and will be described in more detail in the subsequent section on the SDE. It is worth noting, however, that this measure found different patterns of results for the different phases of interest development and found that these patterns changed across different outcome measures. A factor analysis revealed that these questions loaded on the appropriate factors, suggesting that these phases do represent distinct constructs (Wang & Adesope, 2016).

Need for Cognition

Beyond interest in the topic, domain, or that elicited by the text, other predispositions affect comprehension. For example, it may be affected by situational factors (e.g., mood) or more general predispositions (e.g., valuing learning opportunities). As such, these influences are also relevant in understanding comprehension. As the measure used in this study relies on individuals' inference ability, individuals' proclivity toward engaging in complex tasks and going beyond the text is particularly valuable.

Need for cognition (NFC) is a measure of individuals' tendency to engage in cognitively demanding tasks, as well as their enjoyment of these tasks (Cacioppo & Petty, 1982). It has been further suggested to reflect an intrinsic motivation to think and seek information when presented with problems (Furnham & Thorne, 2013). NFC has been used in a wide variety of contexts and is associated with a variety of beneficial effects in education. For example, Meier, Vogl, and

Preckel (2014) found that NFC was the greatest predictor of taking advanced placement courses in a study of German students and Hill et al. (2013) found that NFC was significantly related to general, fluid, and crystallized intelligence measures.

Within the context of reading comprehension, Dai and Wang (2006) suggest that NFC may be expressed through reader's tendency to form a deeper understanding of text, a richer situation model, and to enjoy the act and learning achieved through this deeper engagement. For example, Kardash and Scholes (1996) showed that when individuals with high NFC read texts with conflicting information, they were more likely to accurately represent this mixed evidence than those with low NFC. Additionally, Dai and Wang (2006) found the NFC was a significant predictor of both comprehension and interest measured after reading.

Given the frequent associations of NFC with general learning and comprehension, in particular, it was predicted that NFC would be associated with improved performance on the IIVT.

H_4: There will be a main effect of NFC, with higher NFC scores associated with improved performance on the IIVT.

The Interactive Effects of Prior Knowledge and Interest

As discussed previously, interest and prior knowledge both contribute greatly to student learning. Above and beyond general benefits to learning, different levels of prior knowledge and interest predict what types of materials students will learn the most from (e.g., text difficulty and coherence; McNamara et al., 1996), which types of measures this learning is best captured by (e.g., recall vs. deeper learning; Armand, 2001), and some processes that might facilitate these

improved outcomes (e.g., compensation for poor coherence; McNamara et al., 1996). By examining the interaction between these effects, however, further insight can be gained.

Individual Interest and Prior Knowledge

The relationship between individual interest and prior knowledge has been explored in some studies, albeit much less frequently than studies involving situational or topic interest. Here, the confounding of prior knowledge and interest becomes more acute. That is, individual interest necessarily contains an element of prior knowledge as individuals, by definition, value the opportunity to re-engage with related tasks, generate and seek answers to their own curiosity questions, and are enabled to sustain long-term interest-related tasks, such as creative projects (Hidi & Renninger, 2006). Individuals high in individual interest tend to seek out more information on a topic, leading to greater knowledge. It is worth noting, however, that even this is an assumption and in this extraordinarily sparse literature, there is some support for a more counterintuitive alternate explanation. Specifically, some have argued that it is knowledge that drives interest (e.g., Rotgans & Schmidt, 2017).

A study on attitudes toward science and achievement used measurement of achievement and attitudes at multiple time-points to show that achievement was associated with attitudes toward science, but not the inverse (Mattern & Schau, 2002). Using a repeated-measures design, they found that increases in achievement—for boys, especially—were associated with subsequent increases in positive attitudes toward science. As achievement is related to learning in this context, it can be argued that in learning more, the students became more interested. The presence of a novel gender effect, the nuanced decisions made in statistical analyses (e.g., their criterion for best model-fit), and the use of more complicated analyses limit the confidence one

can have in this single study. However, this study does call into question the common belief that interest leads to increased learning and knowledge.

Recently, Rotgans and Schmidt (2017) conducted a series of experiments to tease apart this relationship. As this research has only appeared in a chapter and was not fully articulated in a peer-reviewed journal, the evidence must also be viewed cautiously; however, it does represent one of the very few studies to look specifically at this phenomenon. In the first study, they measured elementary students' individual interest and knowledge at two time points. They found that interest measured at Time 1 was unrelated to knowledge measured at Time 2, suggesting that interest did not lead to increased knowledge (Rotgans & Schmidt, 2017). However, they found that knowledge at Time 1 did predict subsequent individual interest, suggesting that increased knowledge led to greater interest. Subsequently, they controlled for the effects of prior knowledge on knowledge acquisition and found that individual interest did not affect learning.

Rotgans and Schmidt (2017) then conducted a follow-up experiment on students with high and low prior knowledge and examined interest before and after a learning task. Controlling for individual interest before the learning task, they found that individual interest after learning was still different between those with high and low knowledge, supporting the idea that "interest should be considered a by-product rather than a cause of knowledge gains" (Rotgans & Schmidt, 2017, p. 83). It is important to reiterate, however, that this was a very sparse description of a study, with little information available for more careful examination. For example, it is unclear how individual interest was measured and whether this may instead reflect situational interest which, as described above, makes very different predictions.

Given the paucity of the research on this seemingly critical topic for education, it is unfortunate that these results constitute the majority of evidence for the relationship of individual

interest and prior knowledge. However, they raise an interesting question about whether increasing knowledge may lead to increased interest. As Rotgans and Schmidt (2017) state, "Knowledge changes our view of the world. What appeared previously undifferentiated and bland becomes a source of excitement and opportunity" (p. 83). This idea is certainly intuitively appealing and has some anecdotal support (e.g., increasing interest as people become more competent in their field). Furthermore, they do raise a valid criticism that, although it is commonly accepted that we learn more about the things we are most interested in, it is unclear that true individual interest leads to seeking more information and not that our continued learning elicits an increasingly stronger interest in the subject matter.

This intertwining of the two phenomena has led many researchers to question whether the effects of one can truly be considered without the other. Unfortunately, despite its widespread acknowledgement as a confounding effect, the interactive effects of prior knowledge and individual interest have not received a great deal of attention. The relationship between individual interest and knowledge remains muddy, with few empirical studies. The relationship between topic interest and prior knowledge, however, has received considerably more attention, and has yielded some useful insights.

Topic Interest and Prior Knowledge

Some studies have sought to overcome the challenge of disentangling prior knowledge and interest by studying children, as seeking more information on a topic of interest is more of a product of adult priorities. Baldwin et al. (1985), for example, found little association between the variables, but significant effects of both topic interest and prior knowledge on a recall test in a sample of middle-school students. They explained that as people get older and specialize in certain areas, the relationship between interest and knowledge becomes stronger. Children,

however, are forced to study a variety of topics, regardless of their preference. As such, in these younger samples, it is unsurprising that there would be groups of above-average students with greater knowledge on certain topics (e.g., space exploration) regardless of their topic interest (Baldwin et al., 1985). As children are forced to take a variety of courses, their interest is not as related to their knowledge. This fits well with other research examining prior knowledge and individual interest in children, where the two are not as interrelated (e.g., Rotgans & Schmidt, 2017).

Boscolo and Mason (2003) conducted another study on prior knowledge and topic interest that is particularly relevant to the current project. In this very ambitious study, Italian high school students were given measures of topic interest and topic knowledge (i.e., about the greenhouse effect) which were used to identify students in the highest and lowest quartile, to create four groups, as follows: high knowledge/high interest, low knowledge/low interest, high knowledge/low interest, and low knowledge/high interest. Students read one of three versions of a text about the greenhouse effect that varied in coherence. The minimal coherence text contained minimal word repetitions and connections between topics. The locally coherent text added explicit connections between sentences and used repeated words to increase connectedness, and the globally coherent text included the features of the locally coherent text, as well as topic headers and propositions linking each paragraph to the rest of the text and to the topic as a whole (Boscolo & Mason, 2003).

In addition to the varying levels of topic knowledge and topic interest, students were measured on four different outcome variables, including a free-recall test (i.e., “write down as much of the text as you can remember”) and *bridging-inference questions*, which required participants to connect information from different sentences across the text (Boscolo & Mason,

2003). This task is very similar to the inference measure used in the current study and, thus, provides a close approximation of the expected results. Students were also asked to complete a knowledge-transfer task (e.g., “Is it possible to reduce the emission of gases that increase the greenhouse effect? In what way(s)?”; Boscolo & Mason, 2003, p. 135) and to diagram the greenhouse effect, though the bridging-inference questions are the most relevant for the current project.

Boscolo and Mason found that students in the high knowledge/high interest group performed significantly better on the bridging inference questions than participants in the low knowledge/high interest group or low knowledge/low interest group. Recall performance followed a similar pattern, but with those in the high knowledge/high interest group performing better than those in any of the other three conditions. Regarding text coherence, it was found that individuals in the high knowledge/high interest group performed better with the minimally coherent text than the other two text conditions on both recall and the bridging inference questions. This is interesting as, in general, students learned better from the locally and globally coherent texts, but those with high knowledge and interest learned better from the least coherent text (Boscolo & Mason, 2003).

Although counterintuitive, this aligns well with a study by McNamara et al. (1996), where individuals with high prior knowledge performed best on inference and problem-solving tasks when reading a text that was minimally coherent at the local and macro level. As they suggested, the minimally coherent text led individuals with adequate prior knowledge to attempt to restore coherence to the text (e.g., by generating inferences) which led to improved memory, similar to the more general beneficial effects of more active processing on memory. For example, studies have also demonstrated beneficial effects of *disfluent* texts—texts which are

perceptually difficult to process, like greyscale text or difficult-to-read fonts such as *this example* (e.g., Diemand-Yauman, Oppenheimer, & Vaughan, 2010).

What relevance does this have for the current study? According to the coherence-break hypothesis, also known as the disruption hypothesis, seductive details lead to detrimental learning outcomes by disrupting the coherence of the text (e.g., Lehman, Schraw, McCrudden, & Hartley, 2007). As such, although Boscolo and Mason (2003) only manipulated the coherence of a standard text, without seductive details, their manipulation of coherence may serve as a proxy for the coherence disruption caused by the inclusion of the seductive details. This work may shed light on the findings of Wang and Adesope (2016), who reported beneficial effects of seductive details on transfer performance for those with greater individual interest. As those with higher interest likely have greater knowledge, this could lead to more effective coherence restoration, as proposed by McNamara et al. (1996). This would allow these individuals to more accurately restore coherence, leading to the larger benefit on transfer problems.

The Seductive Details Effect

The Seductive Details Effect (SDE) refers to the finding that the inclusion of highly interesting yet conceptually unimportant information leads to negative learning outcomes, relative to materials without these details (e.g., Garner et al., 1989; Rey, 2012; Wade & Adams, 1990). This literature emerged out of studies examining the role of text-based interest in individuals' learning (e.g., Hidi & Baird, 1986) where the relationship between individuals' interest in the materials and the importance of certain pieces of information for comprehension were considered. It was found that including these highly interesting details actually led to decreased performance.

Studies examining the interestingness of different components and types of text found that text-based interest and importance tended to diverge in certain contexts. Specifically, researchers noted that in narrative texts, importance and interest were related, but in expository texts, they were not (Wade & Adams, 1990). Garner et al. (1989) found that the inclusion of highly interesting, but unimportant details—what they termed seductive details—affected learners' ability to recall the important main ideas of a text, even when instructed to provide only main ideas on a final test. Given the small sample size (n = 36 for 3 conditions), Garner et al. (1989) were cautious in interpreting their results, but speculated that the seductive details may have led the children in their study to inappropriately infer that the main idea of this text was the content of the seductive details (e.g., click beetles), rather than the lives of insects, or that it may be a result of their limited working memory capacity.

Following this work, studies began exploring potential explanations for the SDE (e.g., Schraw & Lehman, 2001), what types of materials could elicit it (e.g., illustrations; Harp & Mayer, 1998), and how individual characteristics may contribute to or mitigate the SDE (e.g., verbal ability; McCrudden & Corkill, 2010). For the purposes of this project, an exhaustive description of all experimental findings will not be provided; however, prominent explanations, critical findings, and their implications are discussed to elucidate the impact of various moderators on the SDE, as well as how they relate to the present work. Despite the contributions of an increasing number of influences on the magnitude of this effect, however, empirical research and meta-analyses (see Sundararajan, 2018 for recent a meta-analysis) have repeatedly supported the idea that the inclusion of seductive details do detrimentally affect the majority of learners.

For this reason, I predict that there will be a significant effect of text condition (i.e., seductive details vs. control), with participants receiving the text with seductive details performing worse on the IIVT than participants receiving the control text without these details.

H_5: There will be a main effect of condition, such that participants who receive the text containing seductive details will perform worse on the IIVT.

Explanations

There are several competing explanations for what causes the SDE (e.g., Harp & Mayer, 1998; Lehman et al., 2007; Magner, Schwonke, Aleven, Popescu, & Renkl, 2014; Rey, 2012; Sanchez & Wiley, 2006). Despite variation in these explanations, however, three overarching theoretical explanations have received the most support, and the most recent versions of each of these theories will be discussed.

Reduced attention hypothesis. The *reduced attention hypothesis*, originally proposed as the *distraction hypothesis* (Harp & Mayer, 1998), suggests that seductive details lead to detrimental effects as individuals' selective attention is drawn toward these details. Harp and Mayer (1998) further suggested that seductive details tend to require little attentional effort to process and understand. This idea was extended by Lehman et al. (2007) to clarify mechanisms and hypothesized influences on reading times and other measures as the reduced attention hypothesis. By their definition, individuals use their finite attentional resources to process the seductive details instead of the base text.

This nuanced revision led to the important prediction that readers would spend less time reading the base text sentences (i.e., sentences not containing seductive details) and that they would remember fewer ideas from the base text. As predicted, Lehman et al. (2007) found that

readers spent less time reading the base text and had poorer recall of the base text. It is worth noting, however, that Harp and Mayer (1998) argued against this explanation by showing that attempts to guide the reader's attention through the use of highlighting, explicit directions to pay attention to the steps of lightning formation, and the use of preview sentences all failed to improve performance in the presence of seductive details. Overwhelmingly, however, other research has supported this explanation (e.g., Chang & Choi, 2014; Rey, 2012; Sanchez & Wiley, 2006), so it is possible that the manipulations used by Harp and Mayer (1998) simply did not effectively overcome the SDE.

Based on the reading-time predictions of Lehman et al. (2007), the current study will examine the amount of time readers spend reading the base text between conditions. Consistent with previous findings, it is anticipated that participants reading the text with seductive details will spend less time reading base text sentences than those who read the control text.

H_6: Participants in the seductive details condition will spend less time reading base text sentences than those receiving the text without seductive details.

Sanchez and Wiley (2006) supported the reduced attention hypothesis by comparing the effects of seductive details on individuals higher and lower in working memory capacity (WMC)—how much information one is able to hold active in memory (e.g., Baddeley, 2003) or, according to the *controlled attention* perspective, one's ability to control their attention in the presence of irrelevant stimuli (Conway & Engle, 1994). Sanchez and Wiley (2006) found that individuals with low-WMC (i.e., lower control of working memory) displayed a typical SDE. Specifically, they performed worse on an essay test and an inference verification task when seductive details were present. Conversely, high-WMC participants displayed marginally better performance ($p < .10$) on the essay and inference verification task when they viewed a text with

seductive details. This finding is in line with the predictions of the reduced attention hypothesis as individuals with high-WMC are better able to maintain focus on the relevant information (i.e., the base text) in the presence of irrelevant stimuli (i.e., seductive details).

Coherence break hypothesis. The *coherence break hypothesis*, a modification of the original *disruption hypothesis* (Harp & Mayer, 1998), suggests that seductive details negatively impact learning by reducing the coherence of the text, leading to a disruption of a coherent mental representation (Lehman et al., 2007). According to the disruption hypothesis, this was the result of interrupting the transition from one main idea to the next in a causal chain of events—in their case, lightning formation (Harp & Mayer, 1998). Lehman et al. (2007) suggested that this was a more general effect of reducing text coherence and the formation of a coherent representation of the text. They hypothesized that individuals would then spend more time attempting to fit the seductive details into this causal sequence and, as such, this hypothesis predicts longer reading times on the transition from the seductive details to the base text (i.e., reading times on the sentence following the seductive details would be greater than the time spent on these sentences in the control text), as well as reduced deeper comprehension. Their findings supported this prediction, with participants in the seductive details condition spending significantly more time reading sentences immediately following seductive details than those in the control condition.

Consistent with this prediction, the current study will also compare reading times on sentences following seductive details and predicts an increase in reading time on these sentences for participants receiving the text with seductive details relative to the control condition.

H_7: Participants in the seductive details condition will spend significantly more time reading sentences following seductive details than control participants reading these same sentences.

Towler and Kraiger (2008) suggested that seductive details are likely to cause reduced recall of a text as it disrupts the *macrostructure* of the text (i.e., the main topic, major subtopics, and relationships between themes). This same disruption, however, has shown positive effects on problem solving tasks as individuals' need to form a general representation of the information being presented leads to increased processing in an attempt to restore this representation, which leads to better performance on more general problem solving (e.g., transfer). This idea is similar to that of disrupting the causal chain of events, and Towler and Kraiger (2008) found that, indeed, participants learning Excel from an audio-visual file with seductive details performed better on transfer than those receiving the training video without such details. More recent studies have also found this pattern (e.g., Wang & Adesope, 2016), though moderating effects were also considered (e.g., interest), so discussion of these findings is reserved for a discussion of moderators on the SDE.

Inappropriate schema hypothesis. According to the *inappropriate schema hypothesis*, originally referred to as the *diversion hypothesis* (Harp & Mayer, 1998), individuals attempt to form their representation of the text around the seductive details, rather than the actual information being learned (Lehman et al., 2007). As Harp and Mayer (1998) explain, regarding their text on lightning formation, students are misled into believing the passage is about "what lightning causes" rather than "what causes lightning" (p. 415). Lehman et al. (2007) suggest that this theory would be supported if individuals recalled more seductive details and had reduced deeper comprehension of the text. Based on the results of their study, the inappropriate schema theory was partially supported in that participants did have reduced comprehension, but they did not recall significantly more seductive details than base text sentences.

Additional evidence for the inappropriate schema hypothesis comes from studies examining the placement of seductive details (e.g., Rowland, Skinner, Davis-Richards, Saudargas, & Robinson, 2008). According to this argument, seductive details placed earlier in texts may preferentially cause learners to form their representation of the text around these seductive details, rather than the base text (Harp & Mayer, 1998). In line with this explanation, some studies have reported larger detrimental effects of seductive details placed at the beginning of a text (e.g., Rowland et al., 2008). Although not significantly different, a recent meta-analysis did find that seductive details placed at the end of learning materials had a weaker detrimental effect (Sundarajan, 2018).

Overview. While all three of these theories are commonly discussed in research on the SDE, it is worth noting that there is some ambiguity regarding their support. Harp and Mayer (1998), who first proposed these explanations in a slightly different form, suggested that it was only the diversion hypothesis (i.e., the inappropriate schema hypothesis) that was supported by their studies. Conversely, subsequent research has given preference to the reduced attention and coherence break hypotheses, as they produce more easily testable predictions (e.g., via reading times) and both have received approximately equal, albeit not uniformly consistent, support. Additionally, according to a meta-analysis by Rey (2012), the inappropriate schema hypothesis has received the least attention and most mixed support.

It is also worth noting that these theories are not necessarily mutually exclusive. For example, an individual could be distracted by the seductive details (i.e., reduced attention), and struggle to restore coherence after this distraction as they have now attempted to incorporate the seductive details into their representation (i.e., inappropriate schema). However, for the purposes of the present research, the coherence break hypothesis can be tested by comparing reading times

on the sentences immediately following the seductive details to the same sentence in the control condition. Additionally, it would be indirectly supported by a finding of improved performance on an inference measure, which resembles a transfer task. Additionally, the reduced attention hypothesis can be partially supported by examining the average reading times for the base text between conditions.

Moderators

As mentioned earlier, there have been several proposed moderators for the SDE including materials-based moderators (e.g., the position of seductive details within a text; Harp & Mayer, 1998), individual differences (e.g., effects of interest; Wang & Adesope, 2016), and task-based differences (e.g., instructions; Peshkam, Mensink, Putnam, & Rapp, 2011). For the purposes of the current study, only materials-based moderators and individual differences will be reviewed, as no task manipulations will be included in the proposed research.

Materials-based moderators. Several differences in the type of materials, or characteristics of the materials, have been suggested to influence the SDE. One example of this is the placement of the seductive details within the text. To test their competing explanations for the SDE, Harp and Mayer (1998) placed their seductive details in several pages at the beginning of a text, interspersed throughout the text, or in several pages at the end of a text. They found that presenting seductive details before the text or interspersed throughout the text both led to worse recall than seductive details placed at the end, which was not significantly different than performance in the control condition. Rowland et al. (2008) also found that seductive details placed at the beginning of a text led to worse performance relative to seductive details placed at the end of a text. The current study placed seductive details either immediately before or after critical sentences—sentences directly necessary for answering IIVT questions—to test this

hypothesis. However, as described in the results section, these data were not analyzed, due to a smaller sample size than initially expected. Pending recruitment of additional subjects, however, it is predicted that seductive details occurring after critical sentences will have a larger detrimental effect on IIVT performance.

Other text-based moderators of the SDE have also been considered, such as the context-dependence of seductive details—whether details are only interesting relative to the rest of the text (e.g., about Microsoft Excel) or are interesting when read in isolation (Ozdemir & Doolittle, 2015)—and the medium the seductive details are presented in (e.g., illustrations; Harp & Mayer, 1998). However, most of these moderators have only been examined infrequently or are not relevant for the current study (e.g., presentation medium). As such, the only materials-based moderator of interest that is relevant for the current study is the positioning of seductive details within the text, which was used to justify interspersing the seductive details throughout the text.

Individual differences. The effects of individual differences on learning from materials with seductive details have received much more attention. Given that the SDE literature emerged out of research on interest—an example of an individual difference—this is not surprising. The most commonly examined individual differences in the SDE literature include various forms of interest (e.g., Magner et al., 2014; Wang & Adesope, 2016) and prior knowledge (e.g., Korbach, et al., 2016; Towler & Kraiger, 2008). Other individual differences have also been examined, such as verbal ability (McCrudden & Corkill, 2010) or working memory (e.g. Sanchez & Wiley, 2006), but much less frequently. For the current study, prior knowledge and interest will be examined, so several key studies are reviewed.

Prior knowledge. The effects of prior knowledge on the SDE has received some attention though, regrettably, it is often included only as a covariate in analyses (i.e., researchers control

for prior knowledge, rather than report it)—echoed in a recent meta-analysis (Sundararajan, 2018). Two studies have provided particularly valuable insight on the moderating role of prior knowledge, however. Korbach et al. (2016) examined the role of prior knowledge and spatial ability on the SDE. Spatial ability was considered because it has been shown to have beneficial effects on education, generally, and because their materials involved the structure and function of ATP Synthase. They found that individuals in the seductive details condition with low prior knowledge performed worse on transfer while individuals with high prior knowledge appeared to be protected from the detrimental effects of seductive details.

Magner et al. (2014) examined the effects of seductive details and prior knowledge on near and far transfer (i.e., different levels of abstraction from the information that students learned). They found that individuals with low prior knowledge displayed a standard SDE on near transfer problems—that is, worse performance for those in the seductive details condition. However, they found that individuals with high prior knowledge performed better on the near transfer problem when learning from a text with seductive details (Figure 1). This beneficial effect of seductive details for those high in prior knowledge seems to resemble the coherence reconstruction effects described by McNamara et al. (1996) and the results of Boscolo and Mason (2003), where individuals with high prior knowledge learned more from texts with reduced coherence, arguably due to increased processing. Interestingly, Magner et al. (2014) included situational interest measures in this study but, as with several other studies, simply used this as a covariate when examining the effects of prior knowledge.

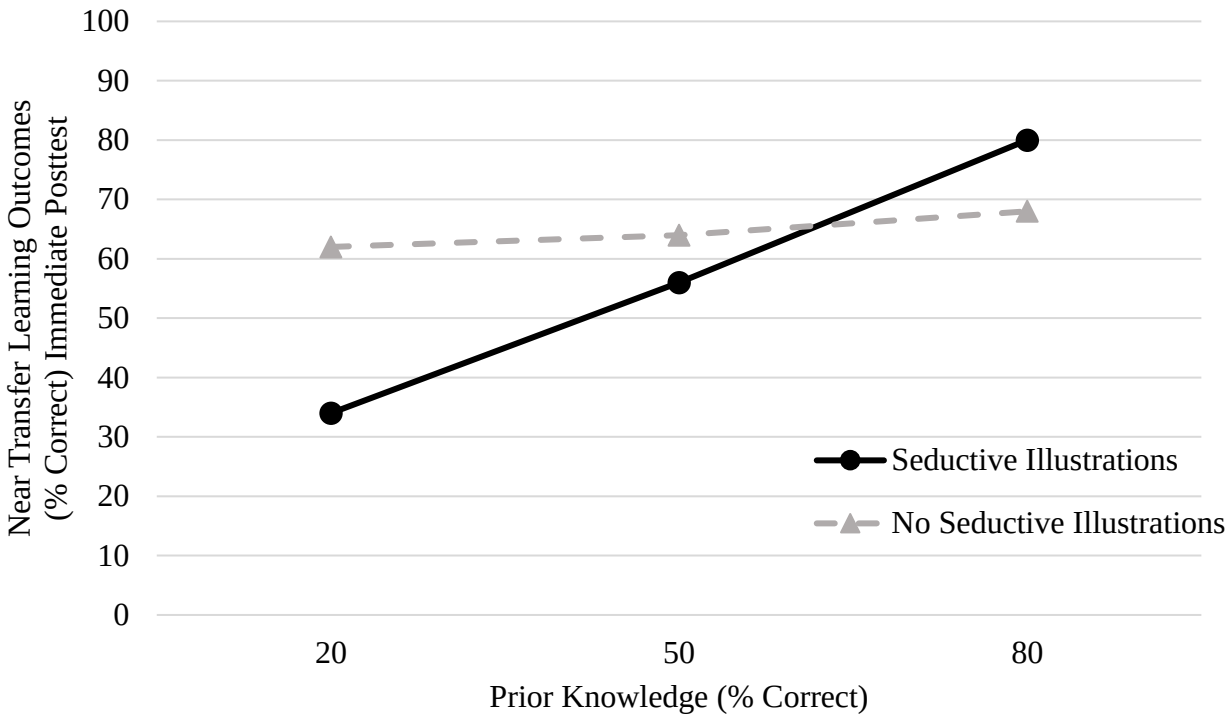

Figure 1. The effects of seductive illustrations on near transfer performance by performance on a prior knowledge test. Adapted from "Triggering Situational Interest by Decorative Illustrations both Fosters and Hinders Learning in Computer-Based Learning Environments" by U. I. E. Magner, R. Schwonke, V. Aleven, O. Popescu, and A. Renkl, 2016, *Learning and Instruction, 29*, p. 149. Copyright 2012 by Elsevier Ltd.

Consistent with Magner et al. (2014), I expect that there will be an interaction between condition and prior knowledge, similar to the pattern seen in Figure 1, such that performance between the seductive details and control conditions will not be significantly different among those with high prior knowledge. Conversely, those with low prior knowledge will show the traditional SDE.

H_8: There will be a significant interaction between prior knowledge and condition. Participants with high prior knowledge will not perform significantly different between the seductive details and control conditions. Conversely, those with low prior knowledge will demonstrate significantly different performance between the control and seductive details conditions.

Interest. As previously described, interest is commonly defined in terms of short-term, usually externally-driven, situational interest; more enduring and knowledge-related individual interest; or topic interest, which may include elements of either or both forms of interest. As the current study will focus on individual interest, in terms of Hidi and Renninger's (2006) Four-Phase Model of Interest Development, Wang and Adesope's (2016) examination of the SDE and the Four-Phase Model of Interest Development provides the most insight.

Wang and Adesope (2016) built upon the SDE literature by examining how seductive details influenced situational interest, as well as how performance was related to individual interest. They developed a scale of 24 seven-point Likert items, from 1 (*strongly disagree*) to 7 (*strongly agree*) to reflect the Four-Phase Model of Interest Development (Hidi & Renninger, 2006). These study materials were presented in Chinese and, although not directly described, it is implied that the scale was initially developed in Chinese, so the English translation has not yet been psychometrically validated. However, based on their analyses, the scale was reliable. Ninth and tenth grade students completed a measure of their triggered situational interest (TSI) and maintained situational interest (MSI) while reading a text on the formation of the Earth. Well-developed individual interest (WDII) was measured immediately before reading and emerging individual interest (EII) was measured immediately after reading the text. Students then completed a multiple-choice recall test, as well as three transfer questions, such as "What would happen to the soil if there were no plants and animals?" (Wang & Adesope, 2016, p. 69).

Wang and Adesope found that seductive details did increase TSI, but that they did not increase any of the other three measures of interest. They also found that seductive details led to decreased recall performance, but there were no differences in transfer performance between conditions. They then conducted mediation analyses to look at the mediating roles of interest on

recall and transfer performance. They found that the detrimental effects of seductive details on recall were mediated by TSI, but not MSI.

Examining the results of individual interest measures, they found that EII did not have a significant interaction effect (i.e., differences in the effects of seductive details on recall for different levels of EII), but they found a marginal effect for transfer performance. Students one standard deviation (SD) below the mean on EII in the seductive details condition performed worse on the transfer test than those in the control condition. At the mean, there was no significant difference and, interestingly, students 1 SD above the mean performed marginally better in the seductive details condition. This pattern, as well as the pattern for WDII, is shown in Figure 2 below. A similar pattern was observed for WDII, where students in the seductive details condition performed worse on recall across WDII. However, as with EII, a reversal of the SDE was found for students with higher WDII. Specifically, students 1 SD below the mean performed worse in the seductive details condition, while no significant differences were found at the mean. Students 1 SD above the mean on WDII performed significantly better on the transfer task. Together, these results suggest that individuals with higher individual interest may perform better on transfer tasks when seductive details are present.

Consistent with the findings and explanations provided by Wang and Adesope (2016), an interaction is predicted such that individuals with higher individual interest will not perform differently on the IIVT, regardless of condition, while a significant SDE will be found among participants with lower individual interest.

H_8: There will be an interaction between individual interest and condition. Participants with higher individual interest will not perform differently on the IIVT between conditions. That is, among those with higher interest, there will be no significant effect of seductive details.

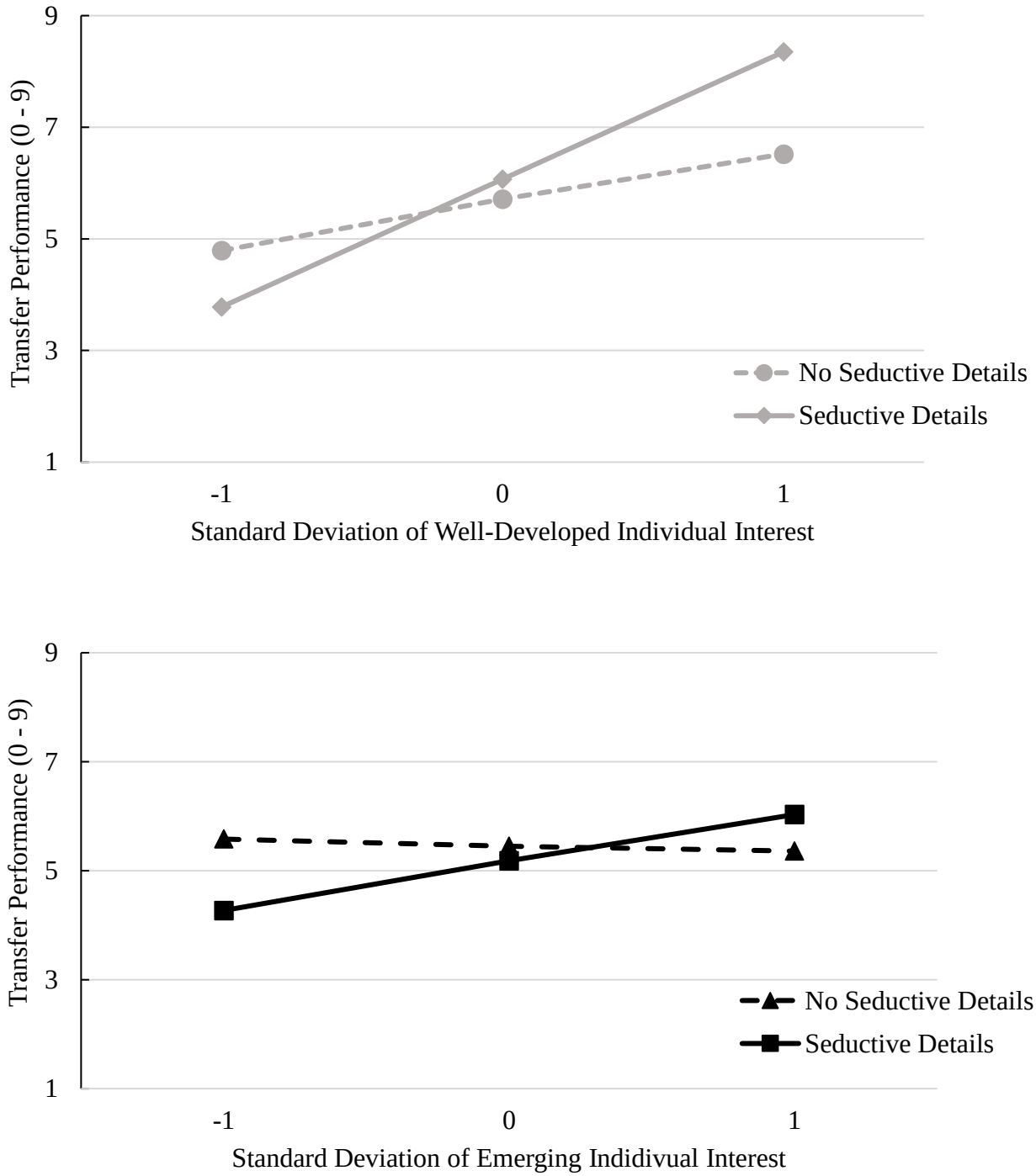

Figure 2. The relationship between well-developed interest and transfer performance (top) and emerging individual interest and transfer performance (bottom). Adapted from "Exploring the Effects of Seductive Details with the 4-Phase Model of Interest" by Z. Wang, and O. Adesope, 2016, *Learning and Motivation, 55*, p. 1424. Copyright 2016 by Elsevier Inc.

This pattern appears counterintuitive at first; however, similar effects have been reported (e.g., Towler & Kraiger, 2008). Given the methodological rigor employed in Wang and Adesope (2016), such as using factor analysis and discriminant function analysis to validate their measure, their effects do not seem spurious. How can this pattern be explained, then? As they suggest, recall measures tend to reflect memory for details of the text, like the textbase level of Kintsch's (1988) construction-integration model, while transfer performance relies on deeper processing and connections to existing knowledge (e.g., the situation model). That is, individuals with higher well-developed interest, who are also likely to have higher prior knowledge (though this was not measured), may engage in deeper processing of the texts with seductive details, thereby improving their deeper understanding (i.e., situation model). This is similar to how texts with reduced coherence lead to readers reconstructing the coherence of the text when they had sufficient prior knowledge (McNamara et al., 1996).

Current Study

Although support for the SDE is quite strong (Rey, 2012; Sundarajan, 2018), recent work has provided new evidence that the inclusion of these details is not uniformly detrimental. These deleterious effects may be moderated by learner characteristics and the type of task involved (e.g., Wang & Adesope, 2016). As such, the current study sought to replicate the SDE and extend it to learning from multiple texts, while considering some of the potential influences on this effect. Specifically, the present study examined individuals' ability to make inferences between texts in the presence or absence of seductive details, while examining the influences of prior knowledge, individual interest (i.e., emerging individual interest and well-developed individual interest), NFC, and vocabulary knowledge.

Pilot Study

As the texts for the current study had not been used previously in seductive details research, a pilot study was conducted to identify the most highly interesting and unimportant of three potential seductive details at each location and to ensure proper functioning of the online presentation of materials via Qualtrics.

Participants and Procedure

A total of 98 participants were recruited through Amazon Mechanical Turk (mTurk), a service which allows *workers* to complete *Human Intelligence Tasks* (HITs)—such as studies, surveys, or other tasks—for payment. To recruit these workers, assignments were posted at 11 staggered times throughout the week in batches of 9 HITs. The first batch was stopped after 8 completions to correct a minor error (an empty question which had been used as a placeholder and still appeared). Workers were directed to a survey created using the Qualtrics survey platform. Upon completing the survey, they were presented with a unique code, generated by a random number feature, which they then entered into the mTurk HIT page. These submitted numbers were then compared to Qualtrics responses (i.e., to ensure the number provided matched a complete Qualtrics response) and a payment was granted. As participation was expected to take approximately 30 minutes, participants were given $4.00 as compensation to reflect a wage of approximately $8.00 per hour.

Materials

Weather texts. Participants read two texts (Appendices A and B, respectively) adapted from Braasch et al. (2013), describing weather patterns across the Pacific Ocean (Text 1) and how these patterns change during an El Niño cycle (Text 2). The first text was 31 sentences (584

words) long and described typical weather patterns across the Pacific Ocean. The second text was somewhat shorter, 12 sentences (255 words) long, and described how weather patterns change during El Niño cycles. For the purposes of the current study, four sentences in the first text, and three in the second, were identified as necessary for answering multiple test questions, hereafter referred to as *critical sentences*, and were used as locations for the seductive details.

Critical sentences were determined based on the information necessary to correctly answer the IIVT questions, described in more detail in the subsequent section. For the IIVT, participants need to combine information from sentences across both texts to make valid inferences, so these sentences were identified and targeted by seductive details. For roughly half of these critical sentences, the seductive detail was placed before the critical sentence, and after the critical sentence for the other half. As previously described, this was done as previous work has shown that seductive details interspersed throughout a text are effective (e.g., Harp & Mayer, 1998); however, their placement within the text (i.e., before or after important information) has not been as thoroughly investigated.

Text ratings. For the pilot study, participants viewed each sentence, one at a time, and rated how interesting each was on a 7-point Likert scale from 1 (*not at all interesting*) to 7 (*very interesting*). They also rated each of these for how important the sentence was for understanding weather patterns on a 7-point Likert scale from 1 (*not at all important*) to 7 (*very important*). For each of the seven critical sentences, participants randomly viewed one of three potential seductive details. As seductive details are defined as highly interesting, yet conceptually unimportant details in a text, the ratings provided a basis to determine which sentence best represented these properties. Below is an example of the base text with three possible seductive details:

In the eastern Pacific, colder waters are pulled up from the depths of the ocean. This process, called upwelling, results in a constant renewal of much cooler surface water temperatures in the eastern Pacific…

[SD1] … When upwelling is disrupted, marine life can be devastated, like in 2012 when warming along the Peruvian coast led to over 800 deceased dolphins washing up on shore.

[SD2] … When upwelling is disrupted, sea life can be devastated, like in 2016 when over 3,000 sea lion pups were stranded in California when warmer waters killed off fish populations.

[SD3] … This area is a primary source of anchovies, despite their controversial status on pizza--one pizza maker reported receiving about 18,000 customers per week at his store, of which an average of only 50 ask for anchovies.

For the location presented above, participants viewed one of the three potential seductive details and rated how interesting and important they thought it was. It is worth noting that although all three examples are tangentially related to the topic of the base text (i.e., upwelling), none of them assist in understanding upwelling.

Results

To determine the most interesting and simultaneously least important seductive detail at each location, participant ratings were collected and average ratings for each potential detail were compared. For several locations, the most interesting potential detail was not simultaneously the least important. For these situations, priority was given to selecting the potential seductive detail with the lowest importance rating as interest ratings tended to be fairly homogenous while importance was much more variable. A difference value was also computed by subtracting the importance from interest ratings to further support these decisions. Ratings of the seductive details chosen for the final study are provided in Table 1.

Table 1

Mean Pilot Ratings of Seductive Details for Current Study

Location		*n*	Interest	Importance	Difference
Text 1					
	SD1	33	5.73	3.30	2.43
	SD2	32	6.00	3.84	2.16
	SD3	33	5.48	3.21	2.27
	SD4	33	5.45	1.91	3.54
Text 2					
	SD1	33	6.12	2.45	3.67
	SD2	33	6.00	2.21	3.79
	SD3	33	5.64	2.70	2.94

Note. Interest and Importance represent mean ratings on 7-point Likert scale. Difference was calculated as Interest - Importance.

Primary Study

Participants

Participants for the primary study were drawn from the university subject pool through the SONA subject pool system. Participants received one credit for participating in the study (equivalent to one hour of participation), which was expected to take between 25-45 minutes, based on timing data from the pilot study. After collecting data for a semester, however, participation rates were low, so additional participants were recruited from two sections of an undergraduate research methods course. Although there were differences in the pattern of reading times as the result of running participants as a group, as opposed to individually, there were no differences in performance between groups. This is described in more detail in the Results section.

Materials

Intertextual Inference Verification Task. The Intertextual Inference Verification Task (IIVT) requires participants to combine information from across two or more texts in order to determine if a provided inference "could reasonably be inferred by combining information from at least two of the texts" and is intended to reflect a measure of deeper understanding of the text than simple recall or recognition tests (Bråten & Strømsø, 2009, p. 15). In the present study, two texts were used and the IIVT was developed in accordance with the guidelines provided by Bråten and Strømsø (2009), such as ensuring that information from both texts is necessary to judge the inference as valid.

An example item states: "During an El Niño event, the trade winds can flow much more slowly from east to west relative to typical weather conditions." To correctly determine that this is a valid inference, participants needed to combine information from Text 1 ("… under normal circumstances, large surface air pressure differences… create the westward blowing trade winds") and Text 2 ("… the more balanced air pressure relationship… causes the trade winds to relax, and sometimes even reverse in direction"). For this study, participants received a total of 18 IIVT questions (7 true, 11 false) and were asked to rate their confidence in each answer on a scale from 1 (*not at all confident*) to 7 (*very confident*). A full list of IIVT items is provided in Appendix C.

Weather texts. The texts for the primary study were the same as those described in the pilot study description, except that seductive details were inserted at the critical locations. Importantly, however, only participants in the seductive details condition read these sentences. Participants in the control condition read the text without these details in order to compare IIVT

performance between conditions, as well as to compare reading times to test various theoretical predictions (e.g., Lehman et al., 2007).

Individual interest. Individual interest was measured using the two individual interest items created by Wang and Adesope (2016). Both individual interest scales are provided in Appendix D. These items were based on Hidi and Renninger's (2006) Four-Phase Model of Interest Development, with several items reflecting each of the four phases. However, for the current study, only the two items reflecting individual interest (i.e., emerging individual interest and well-developed individual interest) were used. Both of these measures were reported to have an interaction effect on transfer performance, with individuals high in well-developed individual interest (WDII) performing better on a transfer task when they received a text with seductive details, as well as a marginal improvement for those higher in emerging individual interest (EII) on transfer performance (Wang & Adesope, 2016).

As the IIVT is intended to measure a deeper representation, I hoped that similar patterns could be seen in the present study, so their interpretation could be aided by additional measurements (i.e., need for cognition, prior knowledge). It is worth noting, again, that Boscolo and Mason (2003) found that individuals with high topic interest and high prior knowledge benefited from a less coherent text on bridging inference questions. A text with seductive details may resemble this, according to the coherence break hypothesis (Lehman et al., 2007).

The WDII measure was included in a prescreening survey for participants in the SONA subject pool—the university's research participation management system. As participants were recruited through SONA, the use of prescreen items allowed us to collect this information early, when participants first set up their account. Thus, there was no inadvertent priming as a result of asking them about their stable, long-term interest in weather immediately before reading the texts

used for this study, as most subjects completed these measures early in the semester and tended not to participate until later. For classroom participants, a Qualtrics version of the WDII and prior knowledge were given a week before participating.

EII was measured after participants had completed the IIVT to assess their growing interest as a result of the information presented in the study, as well as their reflection on the material they read. Although EII did not lead to as large of an interaction for transfer items in Wang and Adesope's (2016) study, it was hoped that this would be related to performance.

Need for cognition. For the current study, the shortened 18-item NFC scale was used, which has been suggested to have similar reliability and is highly correlated with the original 34 item version at $r = .95$ ($p < .001$) (Cacioppo, Petty, & Kao, 1984). For this measure, participants were asked to rate their agreement with 18 questions reflecting attitudes toward various types of cognitive engagement. For example, "The notion of thinking abstractly is appealing to me," or reverse-scored items such as "It's enough for me that something gets the job done; I don't care how or why it works" (Cacioppo et al., 1984, p. 307). A full list of questions for this task is provided in Appendix E.

Vocabulary. In order to measure participants' vocabulary knowledge, and to create a delay between reading the text and answering the IIVT items, a 15-item vocabulary measure was included. For each item, participants received a word (e.g., *belligerent*) and were asked to choose the best definition for the word from five possible options (e.g., *informative, blunt, tiring, war-like* (correct), or *pro-active*) in a multiple-choice format.

Procedure

Participants from SONA came to the laboratory and were first given an informed consent form, describing the nature of the experiment. After agreeing to participate in the study, they were directed to the survey, presented through the Qualtrics survey platform, and read a brief set of instructions. They began by reading both texts, sentence-by-sentence, to collect sentence-level reading times for testing reading time hypotheses.

After reading both texts, participants completed a 15-item multiple-choice vocabulary test. This task was included to serve as a filler task between reading the texts and the IIVT, as well as to measure their vocabulary knowledge, as previous research has shown that this is a strong predictor of inference ability (e.g., Singer et al., 1992). The vocabulary task was chosen as previous research has relied on similar tasks to create a delay between learning and testing to ensure that information is not still in participants' working memory.

After completing the filler task, participants completed the 18-item IIVT with confidence ratings, with instructions presented alongside each inference. Participants then completed the EII, followed by the 18-item NFC scale, and demographic questions (e.g., age, ethnicity). Finally, they read a debriefing form which described the study in more detail, including the purpose of the various parts of the study. A research assistant then asked if they have any questions and dismissed them.

Analyses

Power Analysis

With the increasing emphasis on improving the replicability of psychological science (e.g., Open Science Collaboration, 2015), it has been suggested by many that using a priori

power analysis to determine an appropriate sample size is an important practice, though it is still not done for most studies (Wicherts et al., 2016). Given the limited number of subjects available in the SONA subject pool, the presence of several other in-lab studies, and the much larger number of online studies, a power analysis was conducted, first, based on the assumption that there would be a limited number of subjects available. For this analysis, hereafter referred to as the *simple analysis*, the continuous measures of WDII, EII, and prior knowledge were collapsed into binary groups and NFC and vocabulary were not considered, as the small sample was assumed to lack the power necessary to compare differences between additional variables.

For the second power analysis, the *primary analysis*, all variables were considered in their true form (i.e., as continuous measures) and a larger multiple regression model, including interactions, was considered to explore the relationship between the different individual interest measures, as well as their relations with prior knowledge and NFC. This analysis was the ultimate goal of the project; however, given time constraints, the primary analysis was expected to require additional time for data collection and was intended for publication, rather than for the thesis. As anticipated, participation was too low, so this analysis is not described further.

S**imple analysis.** WDII and EII were first combined into one scale (i.e., the sum of their ratings on both divided by the number of questions). Using a median split (as recommended by DeCoster, Gallucci, & Iselin, 2011), subjects were grouped into a high or low interest group. Similarly, a median split was used to divide prior knowledge into high and low groups. Additionally, interactions between condition, interest, and prior knowledge were planned. A power analysis was conducted in G*Power 3.1. Rey (2012) reported an effect size of $d = .48$ for the effects of seductive details on transfer performance in his meta-analysis, so this was converted to an *f* value ($f = .24$) for use in the power analysis. Consistent with general practices,

an alpha of .05 and power of .80 were selected, and an estimated 139 subjects were suggested in order to examine the main effects and interactions of the simple analysis.

Reading Time Comparisons

Two reading time comparisons were proposed, regardless of the analysis used. That is, even with a smaller sample requiring use of the simple analysis, sample sizes were expected to be more than adequate for these tests. It was anticipated that individual sentence-level reading times would contain some skew. However, while the *t* test does rely on assumptions of normality, it is robust against skewed data, so long as samples are approximately equal in size and sample sizes are fairly large (e.g., 25-30). Further, when these assumptions are violated, it tends to be too conservative (Sawilowsky & Blair, 1992), so does not reflect the more problematic threat of Type I errors.

For the reading time comparisons, first a comparison was made of base text reading time between conditions. As Lehman et al. (2007) argue that seductive details capture the reader's attention at the expense of the base text, the amount of time spent reading base text sentences will be compared. A base text average reading time was calculated for the control condition by adding all sentence times and dividing by the total number of sentences. The same procedure was performed for the seductive details condition except that all times on the seductive details sentences were removed.

A second comparison was made the time spent reading the sentences immediately following the location of the seductive details. For participants in the seductive details condition, reading time for the sentence immediately following the seductive detail at each of the seven locations was recorded. Reading time for these same seven sentences was collected from

participants in the control condition. A series of independent samples *t* tests were then conducted, comparing sentence reading times at each location between the two conditions.

Simple Analysis Results

Due to low rates of SONA participant recruitment, students from two sections of an introductory methods course participated in the experiment. During this time, several behaviors were noted that may have affected the accuracy of reading time (e.g., participants talking to neighbors). As such, a subsequent set of analyses were carried out using more stringent criteria to detect outliers. However, for the sake of transparency, results are first reported for data following the procedures initially proposed.

For this first analysis, two exclusionary criteria were applied. First, as initially proposed, participants with overall survey times 3 Median Absolute Deviation (MAD; Leys, Ley, Bernard, & Licata, 2013) units from the median were removed (n = 2 of 135). Second, inspection of the distribution of predictors showed that some participants had performed below chance on the prior knowledge measure. These participants were removed since this performance suggests that participants were not accurately completing the task. Eight participants were removed from analysis from this procedure (n = 127 total).

Simple Analysis

As indicated previously, simple analysis consisted of 2 x 2 x 2 (Condition [seductive details, control] x Prior Knowledge [high, low] x Interest [high, low]) analysis of variance (ANOVA) with IIVT score serving as the DV. A total interest score was calculated by combining WDII and EII. Total interest and prior knowledge were then split into binary high/low groups based on the median. Descriptive statistics for EII, WDII, and combined interest, as well

as prior knowledge, are reported in Table 2. The three-way between-subjects ANOVA showed a significant effect of Condition, $F(1, 119) = 5.01$, $p = .027$, partial $\eta^2 = .04$, with participants in the Seductive Details condition ($M = 8.71$) performing significantly worse than those in the Control condition ($M = 9.84$). However, contrary to predictions, there were no significant effects of interest, prior knowledge, nor any significant interactions (ps > .10).

Table 2

Descriptive Statistics for Simple Analysis Variables

			Combined		WDII		EII		Prior Knowledge	
Group		*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Interest										
	Low	64	28.19	6.06	8.70	4.37	19.48	4.84	7.44	2.44
	High	63	44.06	5.98	15.27	3.76	28.79	4.39	9.13	2.79
Prior Knowledge										
	Low	72	33.29	10.20	10.57	5.49	22.72	6.69	6.25	1.42
	High	55	39.69	8.47	13.78	4.27	25.91	5.98	10.93	1.50

Note. Combined Interest = WDII + EII; WDII, Well-Developed Individual Interest; EII, Emerging Individual Interest.

Correlational Data

In order to examine the relationship between collected variables, a correlation matrix was created (Table 3). It is interesting to note that both interest measures were significantly correlated ($r = .42$, $p < .0001$) and that prior knowledge was related to both WDII ($r = .24$, $p = .007$) and EII ($r = .22$, $p = .015$), in line with theoretical predictions and previous research.

Table 3

Means, Standard Deviations, and Correlations Using Original Data (n = 127)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. IIVT Score	9.28	2.68										
2. IIVT Confidence	80.72	19.19	.21*									
3. Confidence*Accuracy	4.19	28.71	.96**	.24**								
4. Time on Survey	19.77	5.51	.07	.20*	.06							
5. EII	24.10	6.56	-.03	.38**	-.04	.21*						
6. WDII	11.96	5.23	-.03	.23**	-.02	.22*	.42**					
7. Prior Knowledge	8.28	2.74	.09	.26**	.15	.17	.21*	.24**				
8. Need for Cognition	10.42	13.89	.18*	.38**	.22*	.07	.32**	.14	.27**			
9. Vocabulary	8.77	2.71	.22*	.13	.25**	.12	-.02	.06	.35**	.26**		
10. Avg. Reading Time	7.30	2.92	.04	.08	.04	.85**	.09	.08	.08	.03	.10	
11. Avg. IIVT Time	11.09	4.14	.12	.29**	.11	.79**	.19*	.22*	.12	.15	.10	.60*

Note. EII = Emerging Individual Interest, WDII = Well-Developed Individual Interest, NFC = Need for Cognition. *$p < .05$, **$p < .01$

This provides some convergent support for the validity of these measures, despite none of them being significantly related to IIVT performance (all r's < .10 and p's > .10). Although they were predicted to have weaker relationships with final test performance, IIVT performance was significantly related to both vocabulary knowledge (r = .22, p = .013) and NFC (r = .18, p = .046). As such, these variables were included in subsequent analyses.

It is also worth noting that IIVT performance was not correlated with the amount of time on the survey (p = .424), average reading time (p = .616), or even the amount of time on the IIVT itself (p = .162). As sentence-level reading time data, as well as response time data, are often affected by outliers, a log-transformed version of each time variable was examined. Still, no measures of time were related to IIVT performance. This is somewhat unexpected, as these results suggest that IIVT performance was not related to prior knowledge, interest, or several proxies for effort (i.e., reading time and time on the task).

As confidence was related to many of these measures and reading time seemed to be unrelated to nearly all variables, alternate methods of identifying and omitting outliers were used. These procedures and results are described in the section, Additional Data Cleaning and Results. To foreshadow these results, relationships between timing measures and performance improved following removal of outliers. Additionally, the overall model fit nearly doubled; however, interest and prior knowledge remained non-significant.

Multiple Regression Model

The practice of collapsing continuous variables into binary or ordinal groups is sometimes criticized for reducing power or producing spuriously significant effects (e.g., MacCallum, Zhang, Preacher, & Rucker, 2002), though some have argued against these

criticisms, as well (e.g., Iacobucci, Posavac, Kardes, Schneider, & Popovich, 2015). As such, an alternate analysis was conducted using the original continuous version of these variables. Additionally, performance on the vocabulary test and NFC were included. This analysis also allowed a more direct comparison with data using the additional cleaning procedures.

A multiple regression model was tested to examine the effects of the independent variables vocabulary knowledge, NFC, prior knowledge, and combined individual interest, as well as the binary factor for condition, on IIVT performance. A significant model was found ($F(5, 121) = 3.835$, $p = .003$), with an R^2 of .137 ($R^2_{Adjusted} = .101$). As with results from the ANOVA, the effect of condition was significant, with individuals in the control condition performing significantly better on the IIVT than those in the seductive details condition ($p = .003$). Additionally, both vocabulary knowledge and NFC were significant predictors (p's < .05), while interest and prior knowledge remained nonsignificant. These results are shown in Table 4.

Table 4

Multiple Regression Summary Using Original Data

						95% Confidence Interval	
Variable	*B*	SE_B	β	*t*	*p*	Lower	Upper
Intercept	8.19	1.18		6.96	< .0001	5.865	10.524
Condition	-1.43	0.47	5.10	-3.04	0.003	-2.364	-0.499
Prior Knowledge	0.01	0.09	0.01	0.11	0.911	-0.173	0.194
Combined Interest	-0.01	0.03	-0.04	-0.41	0.682	-0.059	0.039
Vocabulary Knowledge	0.19	0.09	0.19	2.10	0.038	0.011	0.373
Need for Cognition	0.04	0.02	0.20	2.11	0.037	0.002	0.073

Note. 95% Confidence Interval for *B.*

Simple Analysis Discussion

In both the proposed simple analysis—an ANOVA with median split factors for interest and prior knowledge, in addition to the condition factor—and a multiple regression model, prior knowledge and interest were not significant predictors of IIVT performance. While this is contrary to predictions, there are several possible explanations for this pattern of results which are described in the general discussion section. However, both NFC and vocabulary knowledge were significant predictors of IIVT performance. Additionally, the lack of correlation between several online measures (i.e., response time on the IIVT and sentence-level reading times) and IIVT performance warranted further attention.

Additional Cleaning Procedures

Although it was anticipated that MAD would allow removal of participants with suspiciously low times on the survey, low reading times, and low IIVT response times, the original criteria of omitting responses +/- 3 MAD units from the median, as well as the positive skew of these data, did not remove any participants with short times, despite observing some participants "clicking through" the study (i.e., not reading carefully). This was expected as MAD is more robust against skewed data (e.g., Leys, 2013). In order to address this issue and obtain more accurate data for testing reading time hypotheses, additional procedures were used to remove fast responses. Three methods of identifying and removing suspicious responses were implemented, beyond the initial procedures described previously.

First, all timing data were log-transformed to decrease the positive skew of the data. Transformation was applied to all sentence-level reading times, response times, and overall participation time. While this method was very helpful for some distributions, others were still

skewed. Overall, however, transformed data were much more normally distributed. Additionally, while some skew was present, transformation did allow participants with extremely low timing data to be identified.

Second, a set of variables was constructed for each sentence-level reading time, as well as IIVT response times to count the number of responses a participant had that were some number of MAD units from the median. Three sets of each variable were created, reflecting the count of outlier times +/- 2, 2.5, and 3 MAD units from the median. This method was more exploratory in nature, with the goal to identify participants displaying problematically short reading times. However, it is important to note that accuracye on the IIVT, individual differences, or other theoretically motivated measures were not considered when selecting the timing criteria for removal. That is, careful attention was given to removing participants on the basis of problematic patterns of behavior, exclusively, rather than attempting to remove participants to align more closely with theoretical predictions (i.e., p-hacking). Results were consistent regardless of the data considered, with the exception of those changes described explicitly in subsequent sections.

Histograms of these count variables showed that most participants had a very small number of outlier times, while a subset of responses indicated repeated patterns of outlier performance, including one participant with over 40 sentence-level outlier times from a total of 63 possible times. As a variety of methods provided convergent support for the removal of some observations (e.g., the same participants identified for outlier reading times and response times), the strictest criteria of outliers +/- 2 MAD units from the median was used for creating a final cleaned data set. Finally, performance on the vocabulary distractor measure was also used to eliminate participants performing at or below chance (i.e., less than 5). While nearly 90% of participants performed above chance, a subset did not. As such, as previously performed with

prior knowledge, participants scoring five or below (i.e., chance or below) were omitted, removing 10 observations from further analysis.

Final Data Set

As the a priori predicted results included observations known to be problematic (i.e., participants not reading or speaking with neighbors), the strictest cut-off was employed (+/- 2 MAD units from the median) for subsequent analyses. Regardless of whether outliers were dropped based on the number of outlier times on log sentence-level reading times, total number of outlier responses, or response times, the same participants were removed from each procedure. As such, participants were omitted for repeated outlier times on log-transformed reading times, log-transformed average IIVT response times, total number of log-transformed outlier times, or total number of non-transformed outlier times—included to exclude high-end outliers, as well.

These criteria resulted in a much smaller final sample size (n = 97), reflecting the removal of almost 20% of subjects. While this value is quite high, the exclusion of these participants was supported by outlier performance on a combination of criteria. While heuristics often suggest avoiding the removal of more than 5-10% of data, converging evidence suggests that these participants were not participating meaningfully. As such, removing these responses was both quantitatively and logically appropriate.

Cleaned Data Results

Given the large subset of data that was removed, it is important to emphasize that, while defensible, the revised model fitting procedures should be treated as preliminary. While criteria were initially laid out for excluding problematic results, the extent of problematic participation could not be anticipated. As there were increasing issues with recruiting participants through the

subject pool, classroom samples were used which also resulted in some problematic responses due to limitations of running an in-class study (e.g., individuals talking can distract others).

Reading Time Results

Prior to additional model fitting procedures, and in line with a priori predictions, the cleaned data were used to compare sentence-level reading times. Descriptive statistics and results from a series of *t*-tests are shown in Table 5. No significant differences were found, particularly when controlling for multiple comparisons. However, in all but one case, reading times in the seductive details condition were higher than in the control condition, and the average times were consistent with the predicted pattern. As such, this hypothesis was not directly supported, but the pattern of results does suggest that further research using a more homogenous sample may yield the anticipated results.

Table 5

Average Reading Time Differences (in Seconds) on Sentences Immediately Following Seductive Details

	Control (n = 51)		Seductive Details (n = 46)				
Sentence	*M*	*SD*	*M*	*SD*	*t*	*df*	*p*
T1S8	7.53	5.08	9.42	7.14	-1.51	80.4	0.14[a]
T1S17	4.26	2.40	5.10	3.69	-1.32	76.01	0.19[a]
T1S27	11.73	7.54	11.48	6.11	0.17	95.00	0.86
T2S1	6.38	3.10	6.55	3.06	-0.28	95.00	0.78
T2S7	8.46	4.73	8.49	4.88	-0.03	95.00	0.97
T2S12	8.76	5.74	9.78	6.10	-0.85	95.00	0.40
Average	7.85	3.28	8.47	3.23	-0.94	95.00	0.35

Note. [a]Satterthwaite method for unequal variances used. Sentence names reported as T#S## where T = text number and S = sentence number.

Correlations

As a large number of observations were removed, a new correlation matrix was produced to examine the relationship between variables with less influence from suspicious observations. Table 6 shows the correlation between all primary variables of interest. The general pattern of correlations between the cleaned (n = 97; Table 6) and original data (n = 127; Table 3) was consistent, though the relationship between certain variables was notably different.

Relative to the original data, correlations with vocabulary and nearly all variables increased, despite the decreased sample size. A similar pattern is seen between NFC and other variables, with nearly all correlations increasing in magnitude. Several exceptions to this pattern were also seen, however. First, correlations with both interest measures tended to decrease, though the change was generally small (r changes < .10). Additionally, the relationship between timing variables changed. One interpretation of this change is that the removal of primarily low outliers (i.e., participants with exceptionally low times) decreased the tendency for lower times to be associated with lower IIVT performance due to inactive participation (i.e., not reading decreases IIVT accuracy), rather than true reading time. In both the original and the cleaned sample, however, times were largely unrelated to performance (all r's < .15).

Given the generally low association of timing and performance in both samples, correlations were then calculated with observations from the in-class participants removed for the purpose of more accurately analyzing reading times. As individuals in the group setting were much more likely to experience external distractions (e.g., neighbors talking), it was suspected that these data would be less accurate. A revised correlation matrix is shown in Table 7. While the decreased sample size (n = 61) reduced the power of these analyses, many coefficient values increased.

Table 6

Means, Standard Deviations, and Correlations Using Cleaned Data (n = 97)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. IIVT Score	9.38	2.74										
2. IIVT Confidence	80.37	18.26	.22*									
3. Confidence*Accuracy	5.32	29.93	.96**	.28**								
4. Time on Survey	19.83	4.17	.03	.04	-.01							
5. EII	24.16	6.07	-.09	.20*	-.08	.18						
6. WDII	12.21	4.87	-.13	.14	-.10	.17	.35**					
7. Prior Knowledge	8.55	2.74	.03	.28**	.08	.05	.20	.26*				
8. Need for Cognition	11.65	14.03	.19	.45**	.24*	.11	.29**	.11	.21*			
9. Vocabulary	9.14	2.21	.31**	.24*	.35**	.09	.02	.03	.42**	.18		
10. Avg. Reading Time	7.38	2.16	.02	-.06	-.01	.75**	.04	.03	-.09	.04	.03	
11. Avg. IIVT Time	11.18	3.49	.08	.16	.05	.67**	.13	.16	-.01	.21*	.01	.43*

Note. EII = Emerging Individual Interest, WDII = Well-Developed Individual Interest, NFC = Need for Cognition. *$p < .05$, **$p < .01$

Table 7

Means, Standard Deviations, and Correlations Using Cleaned Data with Classroom Observations Removed (n = 61)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1. IIVT Score	9.03	2.66										
2. IIVT Confidence	81.20	16.73	.11									
3. Confidence*Accuracy	1.52	27.15	.97**	.09								
4. Time on Survey	20.38	4.72	.19	.10	.17							
5. EII	24.18	6.02	.02	.35**	.02	.14						
6. WDII	12.23	5.17	-.14	.12	-.12	.17	.24					
7. Prior Knowledge	9.16	2.65	.10	.07	.11	.00	.18	.27*				
8. Need for Cognition	9.41	12.24	.09	.36**	.08	.24	.43**	.08	.19			
9. Vocabulary	8.95	2.00	.37**	.13	.41**	.20	.06	.04	.47**	.12		
10. Avg. Reading Time	7.60	2.18	.23	.01	.21	.74**	.02	.04	-.22	.17	.06	
11. Avg. IIVT Time	11.66	3.62	.25	.21	.23	.76**	.13	.12	-.09	.36**	.09	.51*

Note. EII = Emerging Individual Interest, WDII = Well-Developed Individual Interest, NFC = Need for Cognition. *$p < .05$, **$p < .01$

The relationship between IIVT performance and all timing measures increased substantially (*r* changes > .15), though results were not significant with the reduced power for these analyses. The relationship between most variables and IIVT performance increased, except for confidence and NFC. In general, removing in-class participants led to effects more consistently aligning with expectations. For example, NFC and timing measures were now positively related, reflecting increased effort, and prior knowledge was now negatively related to reading time, likely as a result of increased familiarity and fluency during reading. Overall, patterns were consistent with previous results, but correlations involving timing tended to align more neatly with expectations.

Multiple Regression Model

Using the cleaned data, a multiple regression model was again used to examine the effects of the independent variables, vocabulary knowledge, NFC, prior knowledge, and combined individual interest, as well as the binary factor for condition, on IIVT performance. Despite the decreased sample size, a significant model was found, ($F(5, 91) = 5.36$, $p = .0002$), with a much better model fit value, $R^2 = .228$ ($R^2_{Adjusted} = .185$). As in the original analysis, participants in the seductive details condition did significantly worse on the IIVT than those in the control condition ($p = .004$). Consistent with the original results, prior knowledge and the combined interest measure were not significant predictors (p's > .10); however, both NFC and vocabulary were significant (p's < .05). Table 8 shows the results from this analysis.

Table 8

Multiple Regression Summary Using Cleaned Data

						95% Confidence Interval	
Variable	*B*	SE_B	β	*t*	*p*	Lower	Upper
Intercept	7.61	1.49		5.12	< .0001	4.660	10.568
Condition	-1.55	0.53	-0.28	-2.92	0.004	-2.594	-0.495
Prior Knowledge	-0.11	0.11	-0.11	-0.98	0.328	-0.317	0.107
Combined Interest	-0.03	0.03	-0.11	-1.11	0.272	-0.094	0.027
Vocabulary Knowledge	0.45	0.13	0.36	3.47	0.001	0.191	0.700
Need for Cognition	0.05	0.02	0.24	2.45	0.016	0.009	0.085

Note. 95% Confidence Interval for *B.*

As in the original results, both NFC and vocabulary performance were significant predictors of IIVT performance. Previous studies have found that NFC was associated with increased learning and performance; however, no previous studies have looked at this within the context of seductive details. As NFC is associated with a variety of measures of intelligence (e.g., Hill et al., 2013) and comprehension (e.g., Dai & Wang, 2006), this relationship is consistent with previous research. Vocabulary knowledge is also commonly associated with improved comprehension and learning outcomes, particularly for inferences (Singer et al., 1992), but has also not been examined in the context of seductive details.

Contrary to expectations, and regardless of the selected cleaning method, neither measure of interest were significant predictors of IIVT performance, nor was the combination of these

measures. One possible explanation for this is that the measure asked about their interest in weather, generally. While the texts used in this study were related to weather, they were somewhat dense discussions of El Niño and its relationship with air pressure, which may not align well with those facets of weather that are likely to be viewed as most interesting (e.g., tornadoes, hurricanes). Further, these more interesting facets are also likely more salient.

Another potential reason we did not observe the anticipated effects of interest is that IIVT performance in this study required that individuals remember information from two texts accurately, as well as the source of this information (i.e., to determine if inferences could be drawn between texts). As such, preference for certain topics may be less influential under these more rigorous demands than ability itself—this may also explain why NFC and vocabulary were the only significant predictors of IIVT performance. This interpretation would also explain why the multiple-choice measure of prior knowledge was not associated with IIVT performance, as the recognition-based memory tested by our prior knowledge measure did not reflect the deeper understanding necessary for the current task.

General Discussion

With increasing emphases on active instruction and engaging students—and, perhaps, an aversion to their own dry academic upbringing—many educators have sought to increase the interestingness of instructional materials to improve learning outcomes for students. The current study sought to extend the literature on one potentially problematic approach to making instructional texts more interesting by examining the detrimental effects of seductive details in learning from multiple texts. Results suggest that representations of multiple texts can also be affected, which raises additional questions. For example, could the inclusion of seductive details in a single text lead to detrimental effects of learning from multiple texts? As a great deal of

information is learned by combining information across sources, the current study provides a starting point for examining how learning across texts may be affected by the inclusion of seductive details. Additionally, this study examined the influence of several individual differences on this effect, including learners' individual interest, prior knowledge, vocabulary knowledge, and need for cognition. Participants in the current study were also drawn from a diverse university population, which supports the generalization of these findings to a wider range of learners.

In addition to the inclusion of previously unexplored measures, the current study deviated from the previous seductive details literature in several ways. First, the present study used two texts to explore the effects of learning from multiple texts with seductive details. Second, this study used the IIVT to better understand how seductive details impacted learners' deeper comprehension from multiple texts (e.g., their integrated mental models) using an inference-based measure. Finally, previously-unexamined variables were included in this study (i.e., NFC and vocabulary knowledge).

Seductive Details Effect

As the literature on the SDE continues to grow, it seems increasingly likely that the SDE is a genuine effect. For example, a recent meta-analysis using nearly twice as many studies as the previous analysis by Rey (2012), found consistent support for the SDE, with a moderate overall weighted effect of g = -0.43 (Sundarajan, 2018), which is consistent with the results of this study (g = -0.42). However, instead of questioning the existence of the SDE, current research has focused largely on what boundary conditions exist for the SDE (e.g., Mayer, 2019). For example, recent research has examined the effects of interest (e.g., Wang & Adesope, 2016), prior

knowledge (e.g., Korbach et al., 2016), and working memory capacity (e.g., Sanchez & Wiley, 2006) on the SDE. As such, the current study adds to the literature in several important ways.

Methodological contributions. Several methodological changes were adopted for the current study in order to better understand specific facets of the SDE. Importantly, the current study extended the SDE to multiple texts, incorporated the IIVT for testing learners' deeper representation (e.g., situation model, inference ability), and tested several additional influences on the SDE, such as the placement of seductive details within the text. The use of online timing measures also allowed some insight into the real-time processes that influence the SDE (e.g., reading time, response time). Additionally, in line with suggestions for increased transparency regarding SDE materials (e.g., Sundarajan, 2018), the procedures for creating and modifying the study materials are described.

Validation of study materials. While several trade-offs still had to be made, such as the reduced external validity from presenting the texts one sentence at a time, the current study sought to ensure that all materials were validated by some means to ensure an accurate assessment of the research questions guiding this experiment. In line with the suggestions by Sundarajan (2018) and others, and consistent with the increasing emphasis on transparency in psychological research, procedures for validating these materials were reported in the initial proposal of this research and are summarized briefly here.

First, the texts for the current study were vetted by a domain expert. Additionally, the IIVT items used in the current study were previously created to ensure that the questions referred to information contained across texts (e.g., correct answers could not be produced using just information from one text). This also provided an opportunity to directly target sentences

required for correctly answering IIVT questions by placing seductive details immediately before or after these critical sentences.

The seductive details used in the current study were also validated. Specific results were reported in Table 1. At each of seven intended locations, pilot subjects received one of three potential seductive details and rated how interesting it was, as well as how important it was for understanding the current text. Those sentences that were simultaneously more interesting and less important were selected. Although this method was also subject to limitations (e.g., not all seductive details were the most interesting and least important), we chose the best materials possible for the current study.

Multiple-text comprehension. While many studies of the SDE have examined the effects of seductive details on multimedia learning (e.g., Mayer et al., 2001), the current study is the first to examine the effect of seductive details on learning from multiple text sources pertaining to the same information. Conceptually, learning from multiple texts should be similar to learning from, for example, text and illustrations about lightning formation. However, multiple text comprehension does differ in several ways.

When learners read multiple texts on a topic, the differences between sources of information can be greatly reduced relative to multimedia materials. This is particularly true in the current study, where both texts consisted of plain text on a screen, presented one sentence at a time. This nearly uniform presentation makes sourcing information much more challenging than trying to recall whether the tested information was contained, for example, in a learning text or a supplemental animation. As sourcing information is a critical component of multiple-text comprehension (e.g., Bråten, Braasch, & Salmerón, in press), this makes multiple-text comprehension a particularly challenging task. In the current study, both texts were related to the

same general concept (i.e., weather patterns across the Pacific Ocean), but referred to different components of this—namely, the differences in these patterns during typical and El Niño conditions. As such, while challenging, there were distinct topics contained in each text for learners to use for sourcing information.

Despite the challenges produced by learning from multiple texts, the current study did replicate the seductive details effect. Consistent with previous literature, individuals reading multiple texts with seductive details performed significantly worse on the IIVT than those who read the texts without these details. As such, the current research supports both the existence of the SDE, as well as its generalization to learning from multiple texts.

Intertextual Inference Verification Task. As the current study was the first to explicitly address the effects of seductive details on multiple-text comprehension, it was also the first to use the IIVT. The IIVT is an extension of the IVT intended to capture learners' ability to draw inferences between texts by asking them to judge whether a given inference could be generated by combining information across texts. That is, individuals had to determine if an inference was valid based on information across both texts, as opposed to recognition or another measure which may require only a single piece of information to evaluate.

Testing Explanations

Two primary explanations for the SDE were specifically examined in the current study. A thorough discussion of these theories is included in the Explanations section of this manuscript (pp. 27-32), but they are briefly described here. It is also worth noting that additional explanations and influences on the SDE have been proposed. However, these alternate explanations tend to focus on multimedia learning or do not provide predictions testable by the

current study. Additionally, a third explanation was briefly introduced previously—the inappropriate schema hypothesis—however, this explanation is not discussed further as it did not generate any hypotheses testable by the current study, though it does likely play a role in the SDE.

Reduced attention hypothesis. The reduced attention hypothesis—originally, the distraction hypothesis (Harp & Mayer, 1998)—suggests that seductive details are detrimental as they draw individuals' selective attention toward these interesting, yet unimportant, details at the expense of more relevant and important information. According to Lehman et al. (2007), this explanation would be supported if participants in the seductive details condition spent significantly less time reading the base text (i.e., with time on the seductive details removed) than those in the control condition, as the seductive details should capture a disproportionate amount of the participants' attention.

Results from the current study failed to support this explanation. Using the complete cleaned data, there were no significant differences in average sentence reading time between the seductive details and control conditions on base text. As timing data were considered problematic, this test was also conducted on the data with class observations removed, though results were still not significant. As such, these results do not support this hypothesis.

While the current study does not support the reduced attention hypothesis—nor does it seem likely, assuming a general continuity of the patterns observed thus far—it did not provide many avenues for testing this explanation. For example, Lehman et al. (2007) predicting that, if the reduced attention hypothesis was correct, (1) recall of base text ideas would be worse for participants who received seductive details, (2) that they would spend more time reading the

seductive details relative to the base text, and (3) that participants would recall more idea units from the seductive details than the base text.

As the current study only measured IIVT performance, we could only roughly approximate (1), though this was supported by decreased accuracy among participants who received seductive details. As performed by Lehman et al. (2007), an average reading time per word was computed for seductive details versus control sentences to test (2); however, no significant differences were observed. Finally, (3) could not be tested as the IIVT did not permit intrusions of seductive details into our measure of learning. However, given the low power in the current study and the noisy timing data, these results should be interpreted cautiously. This is particularly true as evidence from previous research using a variety of measures supports this explanation, including relative gaze duration (Chang & Choi, 2014), reading times (Lehman et al., 2007), and WMC (Sanchez & Wiley, 2006). While timing data may not have supported this explanation, future research should also explore online measures to support or refute this hypothesis.

Coherence break hypothesis. According to the coherence break hypothesis—originally, the disruption hypothesis (Harp & Mayer, 1998)—seductive details are detrimental to learning as they disrupt the coherence of the text, leading to a less coherent representation. Lehman et al. (2007) emphasized that this is particularly problematic when seductive details interrupt the transition from one main idea to the next in causal explanations, such as the formation of lightning. This explanation would be supported if participants in the seductive details condition spent significantly longer on base text sentences immediately following seductive details relative to the time spent by control participants on these same questions.

Results from the current study did not wholly support this hypothesis. However, the overall pattern of data was consistent. That is, for six of the seven critical sentences, participants in the seductive details condition had increased reading times, though none of these differences were significant, and most were extremely small: the average sentence reading time for the Control condition (M = 7.85) was less than one second different than that for the Seductive Details condition (M = 8.47). However, given the reduced power for these comparisons and the noisy timing data, it is possible that increased power would support this hypothesis.

Lehman et al. (2007) also predicted that, if the coherence break hypothesis were true, participants who read a text with seductive details should have a worse deeper understanding of the text. That is, while the reduced attention hypothesis proposes a detriment to general recall performance, due to distraction, the coherence break hypothesis posits a particularly detrimental effect of seductive details on deeper measures of understanding. The current study relied solely on the IIVT, so performance between learning measures could not be compared. However, the presence of the SDE in the current study is consistent with this prediction.

Summary. Results from the current study do not strongly support either explanation for the SDE. However, when additional data are collected, it seems plausible that the coherence break hypothesis might be supported, as nearly all effects are in the right direction and several differences are approaching significance. The reduced attention hypothesis should not be considered less likely based on these findings, however, as only reading times could be used to test this hypothesis. Conversely, more precise online measures, such as eye-tracking, have supported this hypothesis.

Learner Characteristics

Individual interest. There were no significant effects of individual interest on IIVT performance, though we initially anticipated a significant effect. This was predicted as interest is often associated with attention, engagement, and learning outcomes (e.g., Hidi, 2001). There was some support for the association of interest with various learning behaviors, as both WDII and EII were associated with increased time on the study and increased average response times on the IIVT; however, these effects were only significant in the original dataset. This does lend some support to the idea that more interested participants may have put more effort into the task, but this was not associated with better performance.

One reason for these findings may be that the IIVT is a much more challenging task than recognition or basic recall performance. This means that, regardless of interest, other factors influence the relationship between interest and performance. The current study is not the first to find that challenging tasks may decrease the association between interest and learning. For example, Gil, Bråten, Vidal-Abarca, and Strømsø (2010) conducted a study examining the influence of writing a summary versus argumentative essay after reading multiple texts. Along with the IIVT, participants completed a sentence verification task (SVT)—a task asking participants whether a statement has the same meaning as a sentence encountered in the text—and the intratextual inference verification task (IntraVT)—a task similar to the IIVT, but testing whether inferences can be made between sentences within the same text. They found that topic interest was related to the SVT (r = .29) and IntraVT (r = .35), but not the IIVT (r = .17). This is interesting as the SVT and IIVT were highly correlated (r = .59), as were the IntraVT and IIVT (r = .58). Together, these results suggest that interest is related to learning, but that this relationship is influenced by task difficulty or the level of comprehension being tested.

Another possible contribution to the non-significant association between interest and IIVT performance may be the more general terminology used in our WDII. This measure asked participants about their interest in weather, broadly. However, the texts used for the current study primarily focused on the relationship between air pressure, precipitation, and weather patterns across the Pacific Ocean during typical and El Niño conditions. Participants may have considered more salient topics within weather while rating their interest, such as tornadoes, volcanoes, or hurricanes. As such, these ratings may not have aligned well with the denser content of the texts that participants encountered.

While the individual interest measures developed by Wang and Adesope (2016) were effective in their study, their modification for this study likely made them too general. Future research should ensure that the scope of the topic measured by these items is consistent with the scope of the learning materials. For example, Clinton and van den Broek (2012) measured topic interest, as well as boredom and engagement, in specific and general topics from the to-be-learned texts and found that interest was moderately associated with inference generation in a coherent text (r = .38). While there were many differences between these studies, their use of general and specific topic interest questions may be a useful approach for future research.

Prior knowledge and interest are proposed to be intimately intertwined (e.g., the affective by-product hypothesis; Rotgans & Schmidt, 2017). As individual interest was only weakly related to prior knowledge in the current study, it is possible that the variability in knowledge and interest was still low, relative to the absolute variability of these characteristics in the student population. That is, high individual interest in our study may still be lower than that of an Earth Science major, for example. Further, in a thorough examination of these relationships on a variety of learning outcomes, Boscolo and Mason (2003) found that *topic* interest—though, their

measured did align well with individual interest—only differentiated performance on inference problems among individuals high in prior knowledge. Additionally, in their study, only participants who scored highest and lowest on interest and prior knowledge were considered (i.e., participants with average values were omitted). As such, if the range of prior knowledge was restricted in our sample, it is possible that our findings may not reflect the general pattern seen in other samples.

Although interest was not a significant predictor of IIVT performance, some interesting patterns were seen. First, EII was significantly associated with participants' confidence on the IIVT. As the EII was administered shortly after the IIVT, this suggests that participants who were more confident in their performance tended to be more interested in the material (r = .35), even after controlling for baseline interest, WDII ($r_{partial}$ = .33). Controlling for actual performance did not decrease this relationship, either ($r_{partial}$ = .35).

Although the current study did not test learning at multiple time points to directly examine changes in interest alongside learning, this relationship suggests that participants' subjective perception of learning—as reflected by confidence in their IIVT performance—predicted their EII, which emphasizes individuals' desire to seek more information on the topic. This finding aligns with Rotgans and Schmidt's (2017) suggestion that individual interest should be considered a by-product of knowledge gains and that perceived learning did increase interest and support motivation to learn more about the topic—though the relationship between EII and actually seeking further information remains untested. This finding also suggests that future research attempting to delineate the relationship between interest and knowledge should incorporate subjective measures of learning, such as confidence ratings or *judgment of learning* measures, alongside actual performance, as true performance may not be as strongly associated

with interest development as subjective perceptions of learning. This may be particularly true when assessments of learning are more challenging or when learners have worse metacognitive accuracy (e.g., young children, less-skilled readers).

Prior knowledge. The current study found no significant effect of prior knowledge on IIVT performance. Furthermore, regardless of the subset of data considered, prior knowledge was only related weakly to WDII and moderately to vocabulary. When classroom observations were removed, it was marginally associated with decreased average reading times ($r = .22$, $p = .08$). Beyond that, however, results were inconsistent between subsets of data, suggesting that these effects may be spurious. As such, even these limited results should be interpreted cautiously.

While many previous studies have measured prior knowledge, relatively few have directly assessed its contribution to learning outcomes in the context of seductive details. Instead, some studies have used prior knowledge as a covariate without reporting additional information (e.g., Jaeger, Velazquez, Dawdanow, & Shipley, 2018) and others measured prior knowledge but never discussed it further (e.g., Sanchez & Wiley, 2006). This is somewhat problematic as prior knowledge is not only relevant for comprehension and memory, but also is associated with individual interest and their interaction has been shown to affect performance, particularly on more challenging tasks (e.g., Boscolo & Mason, 2003). As such, without reporting both of these variables, it is difficult to discern their relative contributions to the influence of seductive details. This sentiment was also echoed in the recent meta-analysis by Sundarajan (2018).

With this limitation in mind, however, many studies that have measured and reported the effect of prior knowledge have found significant effects. For example, in a multimedia study, Korbach et al. (2016) found that the SDE was mediated by relevant picture duration, particularly

for individuals with low prior knowledge. This suggests that individuals with lower prior knowledge spent less time on relevant text and images when seductive illustrations and text were present. This finding provides some support for the distraction hypothesis, as well, by suggesting that individuals spend a disproportionate amount of time on these interesting but unimportant details and that this is particularly true for individuals with low prior knowledge.

Another possible explanation for our non-significant results is the depth of the material covered. As Alexander, Jetton, and Kulikowich (1995) assert, there are distinct, albeit related, forms of knowledge that contribute to interest and recall in learning. *Domain knowledge* refers to a broader, but shallower, understanding of the general domain of knowledge, such as knowing the distinguishing characteristics of psychology versus anthropology. However, knowledge can also include topics within the domain, referred to as *topic knowledge*. While these two forms of knowledge are highly related, they are also separable. For example, one can understand specific topics in psychology—such as Freud's psychosexual stages or genetic contributions to depression—without truly having domain knowledge (e.g., understanding psychology's subject matter). Similarly, one may know generally about anthropology (e.g., the four fields comprising anthropology) without having specialized knowledge about any specific facet of this domain (e.g., cultural relativity).

Our prior knowledge measure likely aligned more with general areas of domain knowledge (e.g., what precipitation is), while the IIVT requires a more coherent and integrated representation of the tested knowledge, such as that contained by individuals with more topic knowledge. As highlighted by Alexander et al. (1995), this is similar to the differences between the relatively fleeting and inconsistent situational interest and the cohesive and integrated individual interest, topic knowledge—which is more common in individuals with limited

exposure (e.g., non-majors)—can be viewed as a more temporary and disconnected set of propositions, rather than the well-structured domain knowledge that may be particularly relevant to inference ability (e.g., the IIVT). Future research should ensure that prior knowledge measures accurately reflect the depth material being tested by the outcome measure.

Need for Cognition. NFC significantly predicted IIVT performance in the current study. While some of the predicted motivational differences were not significant (e.g., individual interest), NFC was consistently associated with participants' performance on the IIVT. It was initially predicted that other variables may have exerted a larger influence on performance (e.g., prior knowledge); however, it is likely that the more difficult nature of the IIVT meant differences in NFC were particularly important.

There are several reasons why NFC was particularly relevant. First, NFC refers to a general intrinsic motivation to engage in, and even enjoy, cognitively demanding tasks. Even in situations where task demands elicit similar cognitive processing among individuals, individuals differ in their enjoyment of these tasks. To clarify, however, while NFC is often associated with improved learning outcomes, it is a motivational characteristic, rather than an actual academic ability (e.g., reading proficiency). As suggested by Cacioppo et al. (1996), it is similar to the differences in individuals' motivation to engage in physical activities (e.g., exercise), which is related to, but distinct from their ability in these tasks. While individuals higher in NFC tend to perform better on various measures of learning—likely influenced by their increased desire to seek out cognitively challenging opportunities—these relationships tend to be modest and distinct (e.g., Fleischhauer et al., 2010).

NFC likely affected individuals' enjoyment of the task and, consequently, the effort they put into their performance. This is supported first by the relationship between EII and NFC. EII

is one of the two individual interest measures used in the current study and reflects individuals' developing individual interest, such as their desire to seek out additional information about the topic. Importantly, EII was administered near the end of the study, shortly after participants completed the IIVT. As such, it measured participants' interest in the topic after being tested. In the complete data, there was a moderate correlation between EII and NFC (r = .32), though this was even stronger when the potentially problematic classroom data were removed (r = .43). Although limited by the post-hoc nature of this comparison, this suggests that individuals higher in NFC were more likely to hold a positive view of the material and task after being tested.

Additionally, when the classroom observations were removed from analysis due to the increased variability of their timing data, NFC was associated with several other indirect measures of motivation and effort. Individuals higher in NFC spent significantly more time on the IIVT (r = .36) and there was a marginally significant association between NFC and overall time on the survey (r = .24, p = .06), despite the much smaller sample size for these analyses (n = 61). Together with EII, these results suggest that individuals higher in NFC were more likely to spend time engaging with the task and that they tended to enjoy it more.

NFC is also related to other behaviors and skills associated with improved comprehension, particularly the deeper forms of comprehension assessed by the IIVT. Dai and Wang (2006) asserted that individuals higher in NFC tend to form a deeper understanding of texts, a more integrated situation model, and to find comprehension tasks more enjoyable. Evans et al. (2003) also found a very strong relationship between NFC and a *deep* approach to learning (from the *Study Process Questionnaire*; Biggs, 1978). Cazan and Indreica (2014) also found that NFC was associated with *deep processing*, *self-regulation strategies*, and several other components of the *Inventory of Learning Styles* (Vermunt & Vermetten, 2004). This frequent

association with deeper processing suggests that NFC provided a particular benefit to participants in the current study, where the IIVT measured a deeper situation model comprehension of the text.

As Cacioppo et al. (1996) suggests, individuals varying in NFC make sense of their world in different ways. Those higher in NFC tend to seek additional information, reflect on their learning, and enjoy cognitively demanding tasks, whereas those lower in NFC tend to adopt more simple heuristics, rely on others' opinions, and adopt a shallower learning strategy. The current study adds to the literature on NFC by showing that individuals high in NFC tend to perform better on a deeper and more cognitively demanding task (i.e., the IIVT), spend more time on these tasks, and the relationship with EII supports their increased enjoyment of these more challenging tasks. Future research should seek to replicate this finding.

Vocabulary. The current study also found that vocabulary knowledge was a significant predictor of IIVT performance. While it was expected that vocabulary would be related to reading comprehension, the strength of this association, particularly relative to other variables in this study, was noteworthy. This finding is consistent with many prior studies on the relationship between vocabulary and inference generation, however (e.g., Singer et al., 1992). While many associations have been reported, no clear mechanism is uniformly accepted. As such, our interpretation of these results avoids reliance on any particular theory and instead reports the similarities between the present work and previous findings.

In one reading comprehension and vocabulary study, Dixon, LeFevre, and Twilley (1988) gave participants three measures of reading, including one measure of inference; three measures of word knowledge, including one vocabulary measure; and three measures of working memory. They found that participants' performance on the Nelson-Denny Reading Test (Brown, Bennett,

& Hannam, 1981) was significantly related to their inference performance (r = .32), as was performance on a number-of-meanings test (r = .32)—a test designed to measure participants' ability to access several word meanings—and their reading span (r = .28). No other measures were significantly related to inference ability. Additionally, in their regression model, only Nelson-Denny scores and the number-of-meanings task were significant predictors of inference performance. These results suggest that vocabulary, broadly measured, is strongly related to participants' ability to draw inferences from text. Interestingly, their inference task also involved reading two texts—though, these texts were narrative, as opposed to the expository texts in the current study—and participants were required to make inferences that could not be answered by relying solely on the information in the text. The inference questions themselves were also intended to measure deeper representation of the situation in the texts. As such, this measure bore some similarities to the IIVT.

Similar results were obtained by Singer et al. (1992) who found a significant relationship between Nelson-Denny Vocabulary scores and far bridging-inference accuracy (r = .31). They suggested that processing more familiar words places fewer demands on participants' limited working memory, which may permit more active manipulation of information and comparisons between word meanings. Further, they argued that while word familiarity may enhance reading span, inferential processes are more likely to be applied to familiar meanings than less familiar ones while controlling for working memory capacity, though this argument was not tested experimentally or statistically. Although short, the texts used in the current study were somewhat challenging (Table 9). As such, the text difficulty may have been particularly problematic for some individuals as the increased effort in decoding word meanings may have required more limited working memory resources.

Table 9

Reading Difficulty Metrics for Texts Used in the Current Study

Text	Flesch Reading Ease	Flesch-Kincaid Grade Level
Text 1	42.3	12.4
Text 2	36.4	14.1

Finally, one previous study did examine the influence of verbal ability on learning from texts with and without seductive details. McCrudden and Corkill (2010) examined recall performance and reading time differences between participants with high and low verbal ability, as measured by a 36-item multiple-choice vocabulary test. In their study, participants were asked to recall as much information from the text as they could. As such, it is a somewhat different measure of memory than the IIVT and reflected a shallower representation of the text. However, consistent with the current study, they found a significant difference in recall performance between individuals with high and low verbal ability.

McCrudden and Corkill (2010) also sought to examine whether verbal ability interacted with seductive details. They found that individuals with higher verbal ability spent less time reading across sentence types. However, regardless of verbal ability, participants spent less time reading base text sentences than seductive details and the longest time reading sentences immediately following seductive details, consistent with the coherence break hypothesis. This suggests that while verbal ability did influence performance, the lack of interaction with sentence type is consistent with the lack of an interaction in the current study.

The significant effect of vocabulary on IIVT performance in our study was consistent with previous research. It is unclear if this benefit was the result of a decreased load on working memory, the relationship between vocabulary knowledge and print exposure, differences in the development of reading skills (e.g., the Matthew effect), or some other facet of comprehension. However, these results are consistent with the literature demonstrating the beneficial relationship between vocabulary and comprehension, particularly for deeper measures of learning.

Limitations & Future Directions

Several limitations affected the current study and are described in detail below. Primarily, these limitations stemmed from the recruitment of participants, though several other limitations were present as a result of trade-offs between external and internal validity (e.g., presentation of text one sentence at a time). Additionally, though decisions were made for testing specific hypotheses, the current study did deviate from previous research on the SDE in several ways, as well as from previous studies on multiple-text comprehension.

Statistical Power

Probably the most apparent and potentially problematic limitation of the current study is the small sample size, particularly after problematic observations were removed. Results from an a priori power analysis suggested that a minimum of 139 participants should be recruited to conduct the Simple Analysis, described in the methods section, which consisted of a three-way ANOVA using a median split to compare combined interest (high/low), prior knowledge (high/low), and condition (seductive details/control). In order to test the full regression model with all participants, an a priori power analysis recommend collecting a total of 246 participants.

Regrettably, issues within the subject pool limited our ability to recruit participants and, after one year of recruitment, we were only able to recruit 139 participants. In order to do so, however, we had to deviate from the proposed recruitment protocol and collect data from a classroom sample, in addition to the subject pool (i.e., SONA) participants. A total of 49 participants were recruited in a classroom setting, while 90 were recruited using from the subject pool. While 139 participants was precisely the number required for the initial analysis, inspection of these data revealed several indicators of potentially problematic responding, including performance below chance on a prior knowledge test (n = 8) and performance below chance on the vocabulary test (n = 10).

Additionally, various issues with outlier times, particularly for sentence-level reading times, were observed. As reading times tend to be more susceptible to errors (e.g., mind wandering), a count variable was created to identify the number of outlier reading times. These procedures are described in detail in the results section; however, to summarize, there were clear patterns within these data, where most participants had relatively few outlier times, while a subset (n = 24) were repeatedly identified as problematic, regardless of the threshold used, and were removed from analysis to ensure that initial results were not influenced by these results. In total, then, only 97 participants were used for confirming the original results, and only data not drawn from classroom samples were suitable for analyses involving reading time.

As a result of the recruitment issues and a pragmatic need to complete this project to meet the requirements of the M.S. degree, these results were reported and interpreted despite their deviation from the proposed procedure and low power. As such, converging evidence was used to support those conclusions that were drawn (e.g., the consistent differences in performance by condition, vocabulary, and NFC). Results should be interpreted cautiously, however, and

additional participant recruitment would be ideal to confirm these results. However, these data do illustrate some informative patterns and suggest some avenues for future research.

Internal Validity

Another limitation was the lack of control over the administration of prescreening measures. For example, eight participants were removed for performing below chance on the prior knowledge measure. Additionally, examination of prescreening data from SONA suggested that some participants were not paying attention. For example, participants who report being more interested in the weather should respond with a lower value on the reverse scored item (*"*weather patterns just do not appeal to me*"*). However, several participants did not do this, suggesting they were not meaningfully participating (e.g., selecting 6 or 7 on all questions).

It is also noteworthy that participants drawn from the classroom sample completed the prescreening measure within two weeks of participating in the study, whereas those who participated through SONA may have completed these measures several months ahead of time. It is unclear how this may have affected performance, but there were some indications that performance was different. For example, the relationship between WDII and EII was much higher among classroom participants ($n = 45$, $r = .59$) relative to SONA participants ($n = 82$, $r = .34$). This is just one example, but it does seem that participants who took the prescreening measure through the Qualtrics survey may have provided more accurate responses.

Reliance on IIVT

While the IIVT captured deeper learning in the current study, it also may have occluded our view of other patterns of learning. For example, neither prior knowledge nor either measure of individual interest were significant predictors of learning. This was somewhat unexpected as

these variables are often associated with learning in the reading comprehension literature, though previous literature has also shown a reduced association between interest and the IIVT (Gil et al., 2010). As such, while these data do not contradict previous findings, they do suggest that our ability to detect some patterns may have been limited. Future research should consider the use of several concurrent comprehension measures, such as the combination used by Gil et al. (2010), to better understand the interactive effects of seductive details and learner characteristics on different levels of comprehension.

Deviations from Multiple Text Literature

It should be noted, as well, that the current study deviated from many studies of multiple-text comprehension in several important ways. First, the present study only used two texts, whereas many other studies employ three or more texts (e.g., Bråten & Strømsø, 2009). While the current study does provide support for a detrimental effect of seductive details on multiple-text comprehension, it would certainly be worthwhile to examine how they may affect comprehension when more texts are used. This may allow researchers to examine the SDE in relation to sourcing of multiple texts, as seductive details may distract learners. Conversely, it may also make each text more distinctive when distinct topics are used, as seductive details are usually tangentially related to the topic of the text.

The current study also differs from some areas of multiple-text comprehension research in that there are few discrepancies between sources of information. As previously research has shown that discrepancies between texts can increase readers' memory for sources of information (e.g., Braasch & Bråten, 2017) it would be interesting to see how seductive details may affect these processes. Similarly, the topic of these texts was not as controversial as those included in many studies, such as climate change (e.g., Bråten & Strømsø, 2009), so it would be interesting

to see how results may vary when reading texts potentially containing belief-inconsistent information which may promote elaboration and increased attention to sources (e.g., through conflict detection).

Deviations from Previous SDE Research

Beyond the use of multiple texts and assessment via the IIVT, the current study also deviated from previous research on the SDE in several ways. First, while many studies on the SDE have used multimedia materials (e.g., seductive illustrations), the current study presented information solely through text. While this does limit how easily these findings generalize to other research on the SDE, Sundarajan's (2018) meta-analysis identified 18 previous studies which exclusively used text and the weighted mean effect size for these studies (g = -0.44) was comparable to those using text and images (k = 25, g = -0.49). These findings are also consistent with our results, g = -0.42. As online courses grow in popularity, lectures use more video and images, and educational materials continue to incorporate more multimedia elements, future research should continue exploring the effects of seductive images and other media; however, previous research has found these effects solely using text, consistent with these findings.

Another area where the current study deviated from previous research is the presentation of texts one sentence at a time. Several previous studies specifically examined reading times for testing competing explanations for the SDE (e.g., Lehman et al., 2007; McCrudden & Corkill, 2010), and these served as the basis for the reading time predictions in the current study. However, presentation of texts in this format has found effects consistent with more traditional presentation methods. As such, while this does limit the ecological validity of the current study, the potential for a better understanding of the online processes underlying the SDE justified this trade-off.

Conclusion

The current study examined the effects of seductive details on participants' ability to judge the validity of inferences across multiple texts, as well as the influence of several learner characteristics on this effect. Consistent with previous research, seductive details led to worse performance, regardless of data cleaning procedures used. Additionally, we found that several individual differences were particularly relevant for moderating this effect, including NFC and vocabulary knowledge. While a larger sample size would be useful to further support these conclusions, this study adds to the literature on the SDE by showing that the detrimental effects of highly interesting, but conceptually unimportant, information applies to inference-based measures and learning from multiple sources of information. Additionally, learner characteristics were particularly relevant, including individuals' proclivity for more demanding tasks and their vocabulary knowledge. Together these findings suggest that educators and individuals involved in the creation of learning materials should be cautious in their attempts to increase the interestingness of materials by the inclusion of fun facts. Instead, alternative methods of increasing interest (e.g., improving cohesion) should be embraced to avoid leaving some learners seduced and confused.

References

Ainley, M. (2017). Interest: Knowns, unknowns, and basic processes. In P. A. O'Keefe, & J. M. Harackiewicz (Eds.), *The science of interest* (pp. 3-24). New York: Springer.

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology*, *94*(3), 545-561.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, *87*(4), 559-575.

Armand, F. (2001). Learning from expository texts: Effects of the interaction of prior knowledge and text structure on responses to different question types. *European Journal of Psychology of Education*, *16*(1), 67-86.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829-839.

Baldwin, R. S., Peleg-Bruckner, Z., & McClintock, A. H. (1985). Effects of topic interest and prior knowledge on reading comprehension. *Reading Research Quarterly*, *20*(4), 497-504.

Biggs, J. B. (1978). Individual and group differences in study processes. *British Journal of Educational Psychology*, *48*, 274-284.

Boscolo, P., & Mason, L. (2003). Topic knowledge, text coherence, and interest: How they interact in learning from instructional texts. *The Journal of Experimental Education*, *71*(2), 126-148.

Braasch, J. L. G., & Bråten, I. (2017). The Discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist, 52,* 167-181.

Bråten, I., Braasch, J. L., & Salmerón, L. (in press). Reading multiple and non-traditional texts: New opportunities and new challenges. In E. B. Moje, P. Afferbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of reading research* (Vol. V). New York: Routeledge.

Bråten, I., & Strømsø, H. I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47*(1), 1-31.

Britt, M., Rouet, J., & Durik, A. (2018). Representations and processes in multiple source use. In J. L. Braasch, I. Bråten, & M. McCrudden (Eds.), *Handbook of multiple source use* (pp. 17-33). New York: Routeledge.

Brown, J. I., Bennett, J. M., & Hanna, G. (1981). *Nelson-Denny Reading Test*. Chicago: Riverside.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116-131.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197-253.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306-307.

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, *114*(*4*), 647-662.

Calvo, M. G. (2004). Relative contribution of vocabulary knowledge and working memory span to elaborative inferences in reading. *Learning and Individual Differences*, *15*, 53-65.

Cavanagh, M. (2011). Students' experiences of active engagement through cooperative learning activities in lectures. *Active Learning in Higher Education*, *12*(1), 23-33.

Cazan, A., & Indreica, S. E. (2014). Need for cognition and approaches to learning among university students. *Procedia-Social and Behavioral Sciences*, *127*, 134-138.

Chang, Y., & Choi, S. (2014). Effects of seductive details evidenced by gaze duration. *Neurobiology of Learning and Memory*, *109*, 131-138.

Clinton, V., & van den Broek, P. (2012). Interest, inferences, and learning from texts. *Learning and Individual Differences*, *22*(6), 650-663.

Cook, A. E., & O'Brien, E. J. (2016). Fundamentals of inferences during reading. *Language and Linguistics Compass*, *11*(7), e12246.

Conway, A. R., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354-373.

Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology*, *137*, 57-75.

Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology*, *32*, 332-347.

DeCoster, J., Gallucci, M., & Iselin, A. M. R. (2011). Best practices for using median splits, artificial categorization, and their continuous alternatives. *Journal of Experimental Psychopathology*, *2*(2), 179-209.

Dewey, J. (2012). *Interest and effort in education*. Miami: HardPress Publishing (Original work published 1913).

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2010). Fortune favors the **bold** (*and the italicized*): Effects of disfluency on educational outcomes. *Cognition*, *118*(1), 111-115.

Dixon, P., LeFevre, J., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*, *80*(4), 465-472.

Evans, C. J., Kirby, J. R., & Fabrigar, L. R. (2003). Approaches to learning, need for cognition, and strategic flexibility among university students. *British Journal of Educational Psychology*, *73*, 507-528.

Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A., & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin*, *36*(1), 82-96.

Furnham, A., & Thorne, J. D. (2013). Need for cognition: Its dimensionality and personality and intelligence correlates. *Journal of Individual Differences*, *34*(4), 230-240.

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of "seductive details" on macroprocessing and microprocessing in adults and children. *Cognition and Instruction*, *6*(1), 41-57.

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Understanding and integrating multiple science texts: Summary tasks are sometimes better than argument tasks. *Reading Psychology*, *30*, 30-68.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193-202.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, *60*(4), 549-571.

Hidi, S., & Baird, W. (1986). Interestingness—A neglected variable in discourse processing. *Cognitive Science*, *10*(2), 179-194.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111-127.

Hill, B. D., Foster, J. D., Elliott, E. M., Shelton, J. T., McCain, J., & Gouvier, W. D. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal of Research in Personality*, *47*, 22-25.

Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015). The median split: Robust, refined, and revived. *Journal of Consumer Psychology*, *25*(4), 690-704.

Jaeger, A. J., Velazquez, M. N., Dawdanow, A., & Shipley, T. F. (2018). Sketching and summarizing to reduce memory for seductive details in science text. *Journal of Educational Psychology*, *110*(7), 899-916.

Kardash, C. M., & Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology*, *88*(2), 260-271.

Kintsch, W. (1980). Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics*, *9*, 87-98.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163-182.

Korbach, A., Brünken, R., & Park, B. (2016). Learner characteristics and information processing in multimedia learning: A moderated mediation of the seductive details effect. *Learning and Individual Differences*, *51*, 59-68.

Krapp, A. (2007). An educational-psychological conceptualisation of interest. *International Journal for Educational and Vocational Guidance*, *7*(1), 5-21.

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, *32*(4), 569-587.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764-766.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*(1), 19-40.

Magner, U. I., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction, 29*, 141-152.

Mattern, N., & Schau, C. (2002). Gender differences in the science attitude-achievement relationships over time among white middle-school students. *Journal of Research in Science Teaching, 39*(4), 324-340.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.

Mayer, R. E. (2019). Taking a new look at seductive details. *Applied Cognitive Psychology, 33*, 139-141.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187-198.

McCrudden, M. T., & Corkill, A. J. (2010). Verbal ability and the processing of scientific text with seductive detail sentences. *Reading Psychology, 31*(3), 282-300.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation, 51,* 297-384.

Meier, E., Vogl, K., & Preckel, F. (2014). Motivational characteristics of students in gifted classes: The pivotal role of need for cognition. *Learning and Individual Differences, 33,* 39-46.

O'Keefe, P. A., Horberg, E. J., & Plante, I. (2017). The multifaceted role of interest in motivation and engagement. In P. A. O'Keefe, & J. M. Harackiewicz (Eds.), *The science of interest* (pp. 49-67). New York: Springer.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.

Ozdemir, D., & Doolittle, P. (2015). Revisiting the seductive details effect in multimedia learning: Context-dependency of seductive details. *Journal of Educational Multimedia and Hypermedia*, *24*(2), 101-119.

Peciña, S., Smith, K. S., & Berridge, K. C. (2006). Hedonic hot spots in the brain. *The Neuroscientist*, *12*(6), 500-511.

Peshkam, A., Mensink, M. C., Putnam, A. L., & Rapp, D. N. (2011). Warning readers to avoid irrelevant information: When being vague might be valuable. *Contemporary Educational Psychology*, *36*(3), 219-231.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, *46*(3), 168-184.

Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and engagement*. New York: Routledge.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, *7*(3), 216-237.

Rotgans, J. I., & Schmidt, H. G. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction*, *32*, 37-50.

Rotgans, J. I., & Schmidt, H. G. (2017). The role of interest in learning: Knowledge acquisition at the intersection of situational and individual interest. In P. A. O'Keefe, & J. M. Harackiewicz (Eds.), *The science of interest* (pp. 69-93). New York: Springer.

Rowland, E., Skinner, C. H., Davis-Richards, K., Saudargas, R., & Robinson, D. H. (2008). An investigation of placement and type of seductive details: The primacy effect of seductive details on text recall. *Research in the Schools*, *15*(2), 80-90.

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, *34*(2), 344-355.

Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and Type II error properties of the *t* test to departures from population normality. *Psychological Bulletin*, *111*(2), 352-360.

Schiefele, U. (1996). Topic interest, text representation, and quality of experience. *Contemporary Educational Psychology*, *21*(1), 3-18.

Schraw, G. (1997). Situational interest in literary text. *Contemporary Educational Psychology*, *22*(4), 436-456.

Schraw, G., Bruning, R., & Svoboda, C. (1995). Sources of situational interest. *Journal of Reading Behavior, 27*(1), 1-17.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review, 13*(1), 23-52.

Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science, 17*(1), 57-60.

Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition, 20*(5), 539-548.

Sundararajan, N. (2018). *Seductive details: A meta-regression and empirical study* (Doctoral Dissertation). Retrieved from ProQuest Dissertations Publishing. (Accession No. 10977803).

Towler, A., & Kraiger, K. (2008). *The effects of seductive details on recognition tests and transfer tasks* (Technical Report No. 1232). Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a483155.pdf

Vermunt, J. D., & Vermetten, Y. J. (2004). Patterns in student learning: Relationships between learning strategies, conceptions of learning and learning orientations. *Educational Psychology Review, 16*(4), 359-384.

Wade, S. E., & Adams, R. B. (1990). Effects of importance and interest on recall of biographical text. *Journal of Reading Behavior, 22*(4), 331-353.

Wade, S. E., Buxton, W. M., & Kelly, M. (1999). Using think-alouds to examine reader-text interest. *Reading Research Quarterly, 34*(2), 194-216.

Wang, Z., & Adesope, O. (2016). Exploring the effects of seductive details with the 4-phase model of interest. *Learning and Motivation*, *55*, 65-77.

Wicherts, J. M., Veldkamp, C. L., Augusteijn, H. E., Bakker, M., Van Aert, R., & Van Assen, M. A. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, *7*, 1832.

Appendix A

Seductive Details (Text One) with Seductive Details Italicized

To understand the reasons that major weather disruptions such as El Niño occur, and why they have such devastating worldwide impacts, one must first consider the forces involved in typical weather patterns. Typical weather patterns in the equatorial Pacific are governed by a strong relationship between the atmosphere and the ocean. Within the Earth's atmosphere, there are high and low air pressure systems that constantly interact with each other. The differences in surface air pressure between the systems cause the wind patterns.

Typically, surface air pressure is higher in the eastern Pacific near the coast of Peru, Chile, and the Galapagos Islands. By comparison, surface air pressure is much lower in the western Pacific near Malaysia, Indonesia, Japan, and parts of Australia. This causes air to move preferentially from the eastern Pacific toward the western Pacific. *Differences in air pressure are also responsible for decompression sickness, or the bends, which causes gas bubbles to form inside the bodies of divers or passengers in unpressurized aircraft cabins and can lead to physical and mental impairments, or even death.* Thus, under normal circumstances, large surface air pressure differences in the equatorial Pacific Ocean create the westward blowing trade winds.

The surface layer of water in the equatorial Pacific Ocean becomes extremely warm. There is a concentrated amount of sunlight at the equator--largely due to the Earth's curvature—which constantly heats the surface water there. As the trade winds steadily push westward against the surface of the ocean, they drag this surface water along. As a result, the warm surface water accumulates in the western Pacific. This bulge of warm water can become quite large, extending out from the coasts of Australia and Indonesia for many thousands of miles. The surface water in the western Pacific becomes the warmest ocean water on the entire planet (75-86º F).

In the eastern Pacific, colder waters are pulled up from the depths of the ocean. This process, called upwelling, results in a constant renewal of much cooler surface water temperatures in the eastern Pacific. *When upwelling is disrupted, sea life can be devastated, like in 2016 when over 3,000 sea lion pups were stranded in California when warmer waters killed off fish populations*. The temperature of the ocean affects the weather that resides directly above it. Because there is a large bulge of warm water accumulating for many thousands of miles in the coastal waters of the western Pacific Ocean, the extremely warm water evaporates faster than colder water. The warmer air temperatures in the western Pacific also facilitate evaporation because warm air can hold more water than cold air. What results is an abundance of hot, moist air above the western Pacific Ocean. The heat also affects these air molecules by making them move around faster and spread out.

As a result, the warm, moist air parcels rise to higher altitudes. At higher altitudes, these air parcels cool down. The cooling reduces the amount of moisture that the air is able to hold, which results in cloud formation. As water accumulates in these clouds, precipitation becomes more likely. This pattern is why countries in the western Pacific, such as Malaysia, Indonesia, and other parts of Southeast Asia have very tropical conditions and experience heavy rainfall. *The heavy rain in the western Pacific, like in Sumatra, allows coffee beans to thrive, though excessive rain (over 120 inches) can lead to fungus problems, like the "Coffee Leaf Rust".*

There is a typical weather pattern that occurs in the eastern Pacific near countries like the Galapagos Islands, as well as Peru and Chile in South America that is quite different than the weather pattern that typically occurs in the western Pacific. Upwelling results in cold surface waters, low evaporation levels, and cooler air temperature in the eastern Pacific. Cooler, drier air is denser, and does not rise to higher altitudes. Altogether, these factors inhibit moisture in the air, resulting in less rain cloud formation and less rainfall. These conditions are why drier, drought-like conditions are the norm in the eastern Pacific, and people in these regions experience very little rainfall by comparison. *Owing to its otherworldly appearance, the Atacama desert in Chile has been used as a location for filming Mars scenes, most notably in the television series Space Odyssey: Voyage to the Planets*.

Appendix B

Seductive Details (Text Two) with Seductive Details Italicized

Every three to seven years, El Niño periods occur, which can last for one to two years. El Niños reflect changes in the physical relationships between the atmosphere and the ocean. But why does El Niño have such drastic effects on weather patterns across the Pacific Ocean and worldwide? Furthermore, what events trigger an El Niño in the first place?

El Niño events occur when the typical surface air pressure differences drastically change across the eastern and western equatorial Pacific. The change in the surface air pressure gradient is due to both lower-than-normal air pressure within the eastern tropical Pacific and a simultaneous higher-than-normal air pressure in the western tropical Pacific. *The 1997-98 El Niño is blamed for an estimated 23,000 deaths and $33 billion in property damage and researchers argued that El Niño may have played a role in a fifth of all civil wars/conflicts from 1950 to 2004*. The more balanced air pressure relationship across the Pacific ultimately causes the trade winds to relax, and sometimes even reverse in direction. Consequently, the massive bulge of warm water--and the warm air that resides above it--rushes towards the central and eastern Pacific. This produces a massive pool of warm water extending in the east for hundreds of miles, which increases the heavy amount of moisture that evaporates into the air.

As more and more moisture evaporates, the hot air rises to higher altitudes, condenses into clouds, and, ultimately, precipitates out as rain. Consequently, areas like Peru and Chile in South America experience heavy rainfall for many months, which often has drastic consequences. *The phrase "raining cats and dogs" originated in 17th century England where, during heavy rainstorms, many homeless animals would drown and float down the street, giving the appearance it had actually rained cats and dogs.*

Conversely, countries in the western Pacific, such as parts of Southeast Asia, experience very little rainfall and drought-like conditions persist during an El Niño event. *Drought linked to a 2007 El Niño sparked a surge in food prices, including an increase in the cost of rice to over $1,000 a ton in 2008 that sparked riots in countries as far afield as Egypt, Cameroon, and Haiti*.

Appendix C

IIVT Questions with Answers in Parentheses

1. During an El Niño event, the trade winds can change direction from eastward to westward. (No)

2. During an El Niño event, air pressure increases in the eastern Pacific Ocean relative to typical weather conditions. (No)

3. Air pressure is higher in the western Pacific Ocean during an El Niño event relative to typical weather conditions. (Yes)

4. During an El Niño event, the trade winds can flow much more slowly from east to west relative to typical weather conditions. (Yes)

5. Westward-flowing trade winds speed up during an El Niño event. (No)

6. The water in the eastern Pacific Ocean is warmer than normal during an El Niño event. (Yes)

7. Upwelling occurs more frequently in the eastern Pacific Ocean during an El Niño event. (No)

8. Upwelling is suppressed in the western Pacific Ocean during an El Niño event. (No)

9. During an El Niño event, evaporation levels are much higher than normal in the western Pacific Ocean. (No)

10. Within the western Pacific, the air is drier than normal during an El Niño event. (Yes)

11. A greater amount of moisture rises to higher altitudes in the eastern Pacific Ocean during El Niños compared to typical weather conditions. (Yes)

12. In the western Pacific Ocean, more moisture rises to higher altitudes during El Niño events compared to typical weather conditions. (No)

13. People living in countries like Malaysia and Indonesia would see more clouds during El Niño compared to typical weather patterns. (No)

14. People living in countries like Peru and Chile would see fewer clouds during El Niño than is typical. (No)

15. Scientists monitoring rainfall patterns in South America would see higher than normal levels during El Niño events. (Yes)

16. Scientists monitoring rainfall patterns in southeast Asia would see higher than normal levels during El Niño events. (No)

17. In countries like Malaysia and Indonesia, El Niño events would change weather patterns from drought-like to more tropical conditions. (No)

18. Countries like Chile and Peru are more likely to experience floods than drought-like conditions during an El Niño weather phase. (Yes)

Appendix D

Individual Interest Measures from Wang and Adesope (2016)

The well-developed individual interest scale was presented as a SONA prescreening measure, and the emerging individual interest scale was presented through Qualtrics, where participants indicate to what extent they will agree with the following statements on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*).

Emerging Individual Interest

1. I will search for more resources about this topic out of school.

2. I am not satisfied with what I learn *only* from this passage.

3. I think I have learned something from this passage that I will remember for a long time.

4. I value this learning opportunity.

5. I would read other articles about this topic if I got a chance.

6. I would make an effort to learn this information again to better understand the content of this passage.

Well-Developed Individual Interest

1. I find learning about weather patterns enjoyable.

2. Weather patterns just does not appeal to me.

3. I value every opportunity to appreciate/learn about weather patterns.

4. It feels effortless when I try to learn something about weather patterns.

5. I think what I learn about weather patterns is worthwhile.

Appendix E

The Need for Cognition Short Form

Items from Cacioppo, Petty, and Kao (1984). This measure was administered through Qualtrics, with instructions for participants to indicate to what extent these statements are characteristic of them on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Items marked with an asterisk denote reverse-scored items.

1. I prefer complex to simple problems.

2. I like to have the responsibility of handling a situation that requires a lot of thinking.

3. Thinking is not my idea of fun.**

4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.**

5. I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something.**

6. I find satisfaction in deliberating hard and for long hours.

7. I only think as hard as I have to.**

8. I prefer to think about small daily projects to long term ones.**

9. I like tasks that require little thought once I've learned them.**

10. The idea of relying on thought to make my way to the top appeals to me.

11. I really enjoy a task that involves coming up with new solutions to problems.

12. Learning new ways to think doesn't excite me very much.**

13. I prefer my life to be filled with puzzles I must solve.

14. The notion of thinking abstractly is appealing to me.

15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.

16. I feel relief rather than satisfaction after completing a task that requires a lot of mental effort.**

17. It's enough for me that something gets the job done; I don't care how or why it works.**

18. I usually end up deliberating about issues even when they do not affect me personally.