Electronic Theses and Dissertations

1-1-2018

The Role of Source Monitoring in Resolving Cognitive Disequilibrium on Texts with Controversial Topics

Shi Feng

Recommended Citation

THE ROLE OF SOURCE MONITORING IN RESOLVING COGNITIVE DISEQUILIBRIUM

ON TEXTS WITH CONTROVERSIAL TOPICS

by

Shi Feng

A Dissertation
Submitted in Partial Fulfillment of the
Requirement for the Degree of
Doctor of Philosophy

Major: Psychology

University of Memphis

August 2018

.

Acknowledgement

My deepest gratitude goes to my mentor and committee chair Professor Arthur C. Graesser, for this persistent patience, consistent support, knowledge, and open-mindedness. Without his guidance and connections I could not have completed my program.

I would like to extend my gratitude to my committee members, Professor Jason Braasch, Professor Gavin Bidelman, and Professor George Deitz, all of whom have taken out their valuable time to contributed resources and genuine support in order to make my journal go smoother than it otherwise would have been. Their valuable suggestions, guidance, and most importantly, diversity of both breadth of knowledge, background, approaches, and ideas enriched me to be a better research scientist.

A special thanks goes to Professor Zhiqiang Cai and Professor Xiangen Hu, two brilliant minded mathematicians whom not only I learned so much from, but they were always willing to help me and listen with care when I was struggling. Without them I could not have accomplished what I have accomplished. Another special thanks goes to Professor Xinhua Yu, without whose expertise in statistics and readily available assistance on writing, this dissertation could not have been possible.

I would also like to thank my CSAL team members: Anne Lippert, Whitney Baer, Qinyu Chen, Genghu Shi, and Ying Fang. Without their help, I could not have finished my program in a timely manner or grow as a researcher.

As always, I thank my family for their everlasting and unconditional support, both financially and emotionally.

And last but not least, I thank my dearest and closest friend, Cristian Moldovan, without whom I could not possibly make it through these difficult years. He understood me in ways that no one else could, and his daily enlightening wisdom, impeccable humor, irreproachable honestly, and penetrating perceptiveness in viewing and assessing life and reality are second to none. I am privileged to have met him and I grew to be a better person than I was because of him. For that, I will forever be grateful.

Abstract

Feng, Shi. Ph.D. The University of Memphis. July 2018. The Role of Source Monitoring in Resolving Cognitive Disequilibrium on Texts with Controversial Topics. Major Professor: Arthur C. Graesser, Ph.D.

This dissertation investigated the effects of text position (i.e. the position that the author of the text who argues in support of a controversial issue), agent roles (tutor versus student), and agent speech acts on source monitoring and how these factors impact memory for information in a learning environment with conversational agents. Previous research on learning from text with conversational agents has supported the claim that agent disagreements stimulate a state of cognitive disequilibrium and confusion in the participants, which leads to deep learning when resolved. Source monitoring is one plausible strategy to resolve conflict engendered from cognitive disequilibrium. This dissertation combines two lines of research (i.e., cognitive disequilibrium and source monitoring) to investigate the interacting effects of agent disagreements with each other, with the texts on controversial topics, and with the participants' personal positions on topics. Participants read texts that had either a pro or con position regarding a controversial topic while listening to a concurrent disagreeing or agreeing discussion between a teacher and student agent regarding the text. The results showed that participants' prior belief was the biggest predictor of memory for text overall. In addition, when participants reported a strong stance regarding supporting or not supporting a topic, the agent disagreement did not affect their discriminatory memory for atypical information. On the other hand, when participants reported a neutral stance regarding supporting or not supporting a topic, the agent disagreements assisted the participants in discriminating the surface structure of the text (i.e., the wording and syntax) as well as their summarization. Source memory was not shown to be a mediating factor for learning during cognitive disequilibrium, whereas the nature of the topics

had a large mediating factor for all memory measures. The results were discussed in the context

of memory models that make predictions on how incongruence with prior beliefs and messages

from texts and agents affects recognition and recall differently.

Table of Contents

List of Tables

List of Figures

The Role of Source Monitoring in Resolving Cognitive Disequilibrium on Texts with Controversial Topics

We are currently living in a digital age where massive amounts of incoming information are available at our fingertips through informational technologies and social media websites. Unfortunately, the reading levels in the United States and other countries have not been keeping pace with the information demands of the modern times. In fact, one in six adults in the United States has been found to have low literacy skills (NRC 2011 Report). There is an urgent need to understand the underlying mechanisms and strategies of deep learning from text. Deep learning is achieved when a reader is able to analyze causal mechanisms, formulate logical explanations, create support for arguments, and resolve conflicts (Graesser, 2015; Graesser, Chipman, Leeming, & Biedenbach, 2009). Deep learning is also associated with better memory retrieval for texts (Brem, Ran, & Pascual-Leone, 2013). As readers learn and encode information into their memory framework, they inevitably encounter contradictory or conflicting information; this is particularly true within digital media, where 62% of Americans receive their news (Matsa & Lu, 2016). The ability to resolve conflicting information from multiple texts on a single topic is critical for gaining an accurate understanding of the topic because it assists in filtering misinformation and developing critical thinking skills (e.g. Jeong, 2003).

A state of mental conflict in face of a new concept that triggers emotion, or cognitive disequilibrium, may occur during learning from texts when learners encounter misconceptions, contradictions, or refutational texts with multiple sources (Kendeou, Braasch, & Braten, 2016; Posner, Strike, Hewson, & Gertzog, 1982). The phenomenon of cognitive disequilibrium has been well documented in the conceptual change and knowledge revision literature where

1

students are presented with new knowledge to actively correct misconceptions (e.g. Kendeou et al., 2016; Kendeou & O'Brien, 2014; Kendeou, Walsh, Smith, & O'Brien, 2014). One consensus from the literature on acquiring new knowledge is the assumption that the learner's prior knowledge or prior belief acts as the reference to which new concepts are being acquired (Limon, 2001). It is ideal to create a *desirable difficulty learning* condition in which the information being learned varies with a learner's prior belief. Interleaving and mixing multiple different topics together can make the learner constantly exposed to new ideas in ways that becomes meaningful. This can be done by triggering the learner's encoding and retrieval processes based on their prior knowledge (Bjork & Bjork, 2014). One way to enhance meaningfulness in inducing conflict is to introduce new concepts or information that would engage the learner on trendy controversial topics. Trendy topics that are controversial often contain conflicting viewpoints because the issues presented are heavily context-relative and no conclusive answers have been offered. These topics will likely result in a learners' cognitive conflict or disequilibrium. From a constructivist perspective (Magoon, 1977), studying the reading of controversial topics may shed light on the various factors that allow a learner to use existing knowledge or prior belief to construct new knowledge from different perspectives. From the experience of the conflict from multiple viewpoints, learners may revise their own belief systems in a form of conceptual or schematic change.

Short-period exposure to prior belief incongruent information is clearly not likely to bring about long-term conceptual change, but it may trigger cognitive disequilibrium in the learner in the short-term. It is this short-term cognitive disequilibrium that is being addressed in this dissertation, not the long-term conceptual change. Although confusion and cognitive disequilibrium have intuitively been viewed as undesirable cognitive affective states during

2

learning, research using computer-generated talking-head agents has shown that deep learning

can occur if the cognitive disequilibrium is accompanied by confusion that ends up being

resolved (D'Mello, Dale, & Graesser, 2012; D'Mello, Lehman, Pekrun, & Graesser, 2015;

Lehman, D'Mello, Strain, Gross, Dobbins, Wallace, Millis, & Graesser, 2011). Cognitive

disequilibrium can be triggered, for example, when someone reads texts on controversial topics

that reflect a lack of congruency between the learner's prior epistemological beliefs and the

actual position of text. Cognitive disequilibrium can also be triggered by listening to an opinion

that contradicts the position of the text and listening to disagreements between two people (or

agents) regarding the position or facts presented in the text (Graesser, Li, & Feng, 2015).

One approach to resolving cognitive disequilibrium is to monitor and consider the source

of information (Braasch, Rouet, Vibert, & Britt, 2012; Braten, Salmeron, & Stromso, 2016;

Goldman, Braasch, Wiley, Graesser, & Brodowinska, 2012; Rouet, 2006). It has long been

established in memory research that the source of information is an important feature that people

use to scrutinize the quality of information. From the scrutiny, one can readily differentiate one

event from another in episodic memory retrieval (Johnson, Hashtroudi, & Lindsay, 1993), but

from the present perspective, more important is how the source of information is relevant to

evaluating the quality of information in various media. When considering the reading of texts,

the source monitoring framework describes how people make decisions about the reliability of

the sources of the texts before, during, or after they read a text. Readers who participate in source

monitoring will encode the information along with its source. During memory retrieval, readers

will evaluate and attribute specific information to its source content and use the source to

facilitate the recall of the information from their memory. According to the *discrepancy-induced*

source comprehension model, discrepancies in text prompt readers to attend more strategically to

sources (Braasch et al., 2012). The connection between the source and its content becomes source-content links in their mental representation of texts (Braasch et al., 2012; Braten, Stromso, & Britt, 2009; Britt & Aglinskas, 2002).

Past research on source monitoring as a process to resolve conflicts during text comprehension has focused primarily on the effect of prior beliefs and the conflicting information presented in the text. However, there are other ways a learner can encounter conflicts when learning from text. For example, in a classroom setting, a teacher or a peer agent can introduce disagreements, either with each other or with the text. The disagreements between agents can induce cognitive disequilibrium in the observer (D'Mello et al., 2015; Lehman et al., 2011). The clashes between the reader's prior belief, text position, agent position, and agent disagreements have yet to be investigated within the framework of individuals experiencing cognitive disequilibrium that will prompt them to use source monitoring as a resolution for conflicts.

Figure 1 depicts a framework for considering different types of cognitive disequilibrium that may occur. The reader's prior belief can clash with the text or with the spoken statements of the agents. Text with controversial topics can clash with the reader's prior beliefs. Agents' spoken statements can clash with the position of the text. Lastly, agents can have spoken disagreeing statements that clash with each other. In a learning environment with both teacher and peer agents, the agents themselves become the sources of information. An investigation of both the learner's memory of the text and the agent sources should shed light on how the learner achieves deep learning when cognitive disequilibrium is present.

The goal of this dissertation is to investigate the memory of texts and spoken messages after participants read controversial texts with different positions using *AutoTutor,* an artificially

intelligent learning environment with agents (Graesser, 2016; Graesser, Cai, Baer, Olney, Hu, Reed, & Greenberg, 2016). Potential cognitive disequilibrium and confusion were manipulated in two ways. The first condition was by having the participants read texts that either contradict, or do not contradict, with their prior beliefs about a position on a controversial issue. The second condition was by having the agents disagree with each other and the text, compared to an agreement between the agents. The current study collected summary scores, agent source memory scores, and recognition memory for sentence surface structure and meaning. Source memory scores was also used as a predictor of the summary scores and sentence recognition memory scores. The dissertation investigated alternative models and hypotheses that relate to cognitive disequilibrium, source memory, recognition memory for sentences, and comprehension for text.
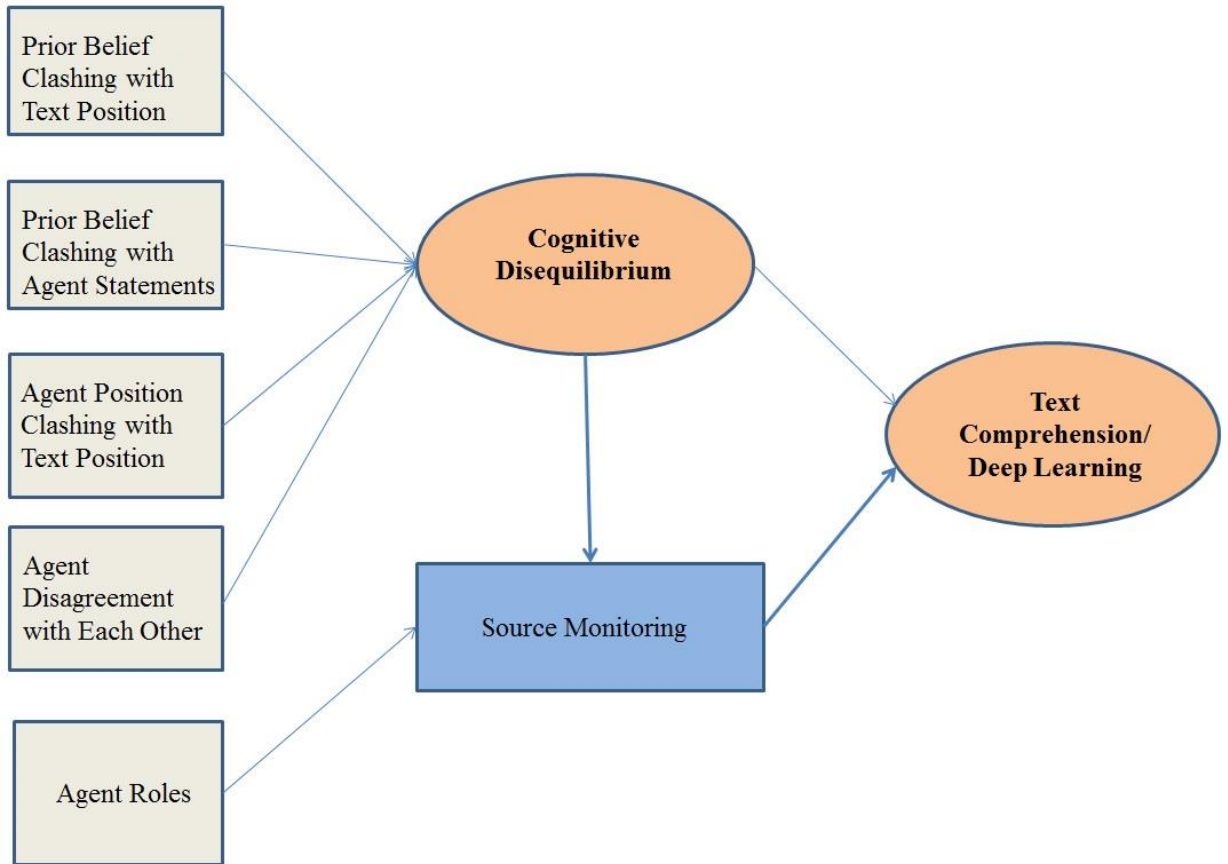
Figure 1. Theoretical framework that underlies the current study.

Mechanisms of Cognitive Disequilibrium in Learning from Text

Cognitive disequilibrium during reading comprehension occurs when the learner is unable to successfully construct a coherent situation model of the text because of an occurrence of contradictory information either with the learner's prior knowledge schemas or contradictory information within the text itself (Graesser et al., 2015). The *situation model* is a mental representation of the text that is constructed and updated as the reader actively comprehends and integrates incoming information (Graesser, Singer, & Trabasso, 1994; Kintsch, 1998; Zwaan & Radvansky, 1998). A break in coherence from contradictory information puts the reader in the state of cognitive disequilibrium. If the situation in the text contradicts the learner's prior beliefs, the learner also may not successfully construct a coherent situation model, or alternatively will compartmentalize a coherently constructed message as unreliable or incorrect. Previous research has reported that this cognitive disequilibrium may produce deeper learning (D'Mello et al., 2015). For example, when readers encounter a situation in the text that contradicts earlier information or their prior knowledge, proficient readers spend more time reasoning and generating more inferences, which results in a deeper learning as reflected in memory measures (Blanc, Kendeou, van den Broek, & Brouillet, 2008; Cook, Halleran, & O'Brien, 1998; Maier & Richter, 2013; O'Brien, Rizzella, Albrecht, & Halleran, 1998; Rapp, 2008). If readers manage to resolve the contradiction, they presumably will achieve deeper learning, but it is also possible that a resolution is not necessary; deep learning may possibly be achieved by virtue of the inferential reasoning itself, even without a successful resolution.

Creating the conditions under which the agents disagree over the content of the text is another way to induce confusion and promote deep learning. Recent research has supported the view that disagreements in conversations between an instructor agent and a peer agent prompt

the learner to evaluate the veracity of the claims from both sides and results in deeper learning

(D'Mello et al., 2015; Graesser et al., 2015; Lehman et al., 2011; Lehman et al, 2014). These

studies used *AutoTutor trialogues* (Cai, Feng, Baer, & Graesser, 2014; Cai, Graesser, Forsyth,

Burkett, Millis, Wallace, Halpern, & Butler, 2011). Participants were placed in a learning

environment with two talking head agents. One agent assumed the role of the instructor, and the

other agent assumed the role of the peer. They read case studies and were asked to read and

critique the studies using their knowledge of research methods in psychology. In a True-True

control condition, the teacher and the student agent agreed upon the correct critique. In a True-

False condition, the teacher gave the correct critique, whereas the student disagreed and

expressed an incorrect assertion. In a False-True condition, the teacher gave the false assertion

whereas the student disagreed and expressed a correct assertion. Finally, in a False-False

condition, both the teacher agent and the student agent agreed upon a false assertion. D'Mello et

al. (2015) found that participants reported more confusion during the conditions when the agents

disagreed with one another (i.e. True-False and False-True). That is, performance scores on

multiple-choice post-tests and transfer tests used to assess deep learning were higher when the

agents disagreed with each other when compared to the control True-True condition. Moreover,

confusion on the part of the participants had a moderating role in deep learning. That is,

participants who did not experience confusion did not benefit from the manipulated cognitive

disequilibrium as much as the participants who did experience confusion. This finding supports

the hypothesis that confusion plays a moderator role in deep learning. When a learner is

confused, resolving the confusion leads to deep learning, but it is an open question whether a

resolution is an important prerequisite to deep learning. Perhaps confusion plus deliberation

without resolution is sufficient for deep learning.

Consider an AutoTutor trialogue when the learner is reading about a controversial topic. The learner may disagree with both the position of the text he/she has read and with at least one of the agents. If the teacher agent agrees with the position of the text—while the learner disagrees with it—then the learner will also disagree with the teacher agent. This relationship is interesting to investigate because the instructor is often appraised as a reliable source of information that the learner trusts (Rourke & Anderson, 2002). As another possibility in the trialogues associated with a text, the learner may encounter a peer agent who agrees or disagrees with the position of the text, which in turn can clash with the learner's prior beliefs. While having agents disagree with one another will put the learner into a state of cognitive disequilibrium, the agents' position that contradicts the position of the text may increase the confusion of the learner even more. Thus, the learner must incorporate reading strategies to resolve the cognitive disequilibrium or otherwise run the risk of failing to comprehend the text.

Source Monitoring and Source Memory in Comprehension

When learners are confronted with opposing viewpoints from reading more than one text on the same topic, the source information of each of the texts can assist the learner in resolving any conflict they may encounter. The source of an article or text is an informational artifact that can refer to the author or a publication site (Lawless, Goldman, Gomez, Manning & Braasch, 2012; Wiley, Goldman, Graesser, Sanchez, Ash, & Hemmerich, 2009). Actively monitoring the source information, called *source monitoring*, is a strategy for scrutinizing the quality of information and making decisions regarding the information to which attentional resources are better allocated. Memory for the source can also help retrieve text information to be used for resolving any confusion the learner may have while reading texts with contradictory opinions or viewpoints (Lawless et al., 2012; Rouet, 2006; Wiley et al., 2009).

The ability to distinguish between reliable and unreliable sources should help a reader determine which authority to attend to. A deliberate strategy to make such distinctions is useful in resolving conflicts between texts with different positions (Goldman, et al., 2012; Goldman, Ozuru, Braasch, Manning, Lawless, Gomez, 2011; Stadtler, Scharrer, Brummenhenrich, & Bromme, 2013). According to the *document model framework*, when the learner recognizes whether the previous information from one document contradicts the current information that has been read, links between the source and the content are formed (Britt & Rouet, 2012). This link then gets encoded into the reader's memory framework. Goldman et al. (2012) conducted a study which found that better learners, when compared with poor learners, monitor source information more and utilize sources more effectively during reading. Skilled readers are also better at integrating source information into their memory of the situation model (Goldman et al., 2012; Goldman et al., 2011). Similarly, Anmarkrud, Braten, and Stromso (2014) used think-aloud procedures to investigate how students strategically read multiple-documents about an unfamiliar scientific issue while evaluating the trustworthiness of document sources. They found that students who primarily used evaluating as a reading strategy tend to cite or use the source or attempt to evaluate the trustworthiness of the sources. In addition, the more that the evaluation strategy was utilized, the more the students favored sources with a comparatively balanced view over a source with a more biased view. Hence, encoding sources into the reader's memory framework assists the reader to make better decisions when they are faced with conflicting information during multiple-document comprehension.

It is plausible that the source monitoring strategy is initiated under specific conditions rather than uniformly. According to the *discrepancy-induced source comprehension model*, when readers encounter conflicting accounts of a situation or opposing points of view, they increase

their attention to sources during encoding, so source memory is enhanced in a subsequent memory test (Braasch et al., 2012; Braten, Stromso, & Britt, 2009; Britt & Aglinskas, 2002). For example, Braasch et al. (2012) examined participants reading critical stories describing news events (i.e. topics extracted from the internet, which included science, society, and the economy). The stories were created either with or without discrepancies. They found that participants allocated more attention to sources when encountering stories with discrepancies and were better able to recall the sources of the discrepant stories than those of stories without discrepancies (Braasch et al., 2012). This suggests that encountering conflicting rather than agreeable information would encourage a learner to actively monitor the source of that information.

Previous research has also confirmed that learners attend more to the source of the text when the text information contradicts the learner's prior topic beliefs. Braten et al. (2016) asked participants to read one of two versions of a text that discuss the controversial topic of whether cell phone radiation has potential health risks. They found that the participants' memory for source information was best when their prior belief was that cell phones posed health risks, and they read the text that cell phones posed no risks. Braten et al. (2016) explained this finding based on the *Plausibility-Induced Source Focusing Assumption*, which states that when different texts make conflicting claims about a controversial issue, linking the discrepant content information to the sources of the texts can help resolve the break in coherence of the mental representation of the situation discussed between the texts (de Pereyra, Britt, Braasch, & Rouet, 2014). Interestingly, when the participants held the prior belief that cell phones posed no health risks, there was no significant difference in source memory between the text that was congruent with their prior beliefs and the text that contradicted those beliefs. This suggests that additional factors may influence a learner's memory for sources.

Memory for Text in Reading Comprehension

Memory measures for the text are one useful method to assess the effect cognitive disequilibrium has on learning. Text memory often predicts comprehension (Daneman & Merikle, 1996). Past research has shown that different memory measures, such as recognition versus recall, result in different memory saliency when encountering prior belief congruent and incongruent information (Stangor & McMillian, 1992). One classic model, the Semantic Associative Memory model (SAM; Raaijmakers & Shiffrin, 1981), suggests that memory information contains item, associative, and context. Retrieval from long term memory is based on context clues and whether that information is retrievable would be dependent on the associative strengths of the cues to that memory. This means that in recall tasks, the search process involves a series of attempts based on the relative retrieval activation strengths. Therefore, memory recall for an item would be facilitated when the learner's prior knowledge or prior beliefs have stronger association with that item.

However, there is no guarantee that semantic similarity of concepts could also facilitate discriminatory memory. The SAM model emphasized that memory recognition, contrary to memory recall, does not necessarily involve sequential attempts but can be from a single-step retrieval step based on activation strengths (Raaijmakers & Shiffrin, 1981). Previous research has found that discrimination is better for unrelated concepts or atypical information (Begg, 1978; Epstein, Phillips, & Johnson, 1975; Hunt & McDaniel, 1993). Discrimination is described as separating a target signal from noise. In the case of assessing learning using memory measures, discrimination typically involves identifying a correct old item when provided with multiple alternatives. Hunt and McDaniel (1993) predict that distinctive or atypical items are more likely

to be distinguished in tasks that rely on recognition rather than information generation in memory recall.

Past research has suggested that readers tend to remember information from the text better when the text is consistent with the reader's prior belief and general knowledge (Stangor & McMillan, 1992). This makes it easier for the reader to construct a coherent situation model which will enhance comprehension and memory for text (Kintsch 1998; Voss & Wiley, 2000; Wiley & Voss, 1999). On the other hand, memory frameworks are dependent upon the individual's schemas, or generic knowledge structures, that guide the reader's interpretations, inferences, expectations, and attention (Graesser & Nakamura, 1982). During text comprehension, the *Schema Pointer + Tag* hypothesis states that the memory trace consists of a pointer to the generic schema and a set of tags for information that is atypical of the schema (Graesser & Nakamura, 1982). Information that is atypical would include information that contradicts the reader's prior beliefs and expectations. From this perspective, with regards to schemas, readers tend to have better memory discrimination for atypical information that violates their schemas and, therefore, remember it better. In support of the Schema Pointer + Tag hypothesis, past research has found that when participants were presented with stories containing scripted activities, such as eating in a restaurant, memory discrimination was virtually absent for very typical script action, whereas it increased in accuracy as typicality of the action decreased (Graesser, Woll, Kowalski, & Smith, 1980). Furthermore, it has found that atypical scripted actions had faster forgetting rate than typical actions in memory recall, whereas for memory recognition, the forgetting rate was the same for both atypical and typical actions (Graesser et al., 1980; Smith & Graesser, 1981). From these classical memory models, the effect of cognitive disequilibrium on learning could be investigated with respect to participants' prior belief

congruence to position of the text. That is, prior belief that is compatible with the position of the text should have greater retention for information generation in memory recall, whereas prior belief that is incompatible with the position of the text should have better memory discrimination, or recognition.

One method to assess memory recall as well as deep comprehension is to ask the reader to write a summary of the text they have just read. Writing a summary of the text requires the reader to recall information from the text and formulate the information in a coherent (albeit condensed and simplified) manner (Brown, Bransford, Ferrara, & Campione, 1983). A good summary is comprised of the most germane information from the text. The reader must be able to discern and self-explain main ideas and supporting information of the ideas. The reader must also be able to actively identify and ignore irrelevant information in order to integrate the central main ideas in a meaningful way (Brown, et al., 1983; Brown & Day, 1983; Dascalu, Stavarache, Dessus, Trausan-Matu, McNamara, & Bianco, 2015). The quality of summary writing would, therefore, provide both memory recall measures as well as the reader's overall comprehension level. It might also make predictions regarding prior belief congruence based on the previous memory research on information generation (e.g. Hunt & McDaniel, 1993). Simply put, a summary task should benefit typical congruent information over atypical incongruent information. Consequently, summary writing is an integral phase of the current study.

A method to assess the reader's memory discrimination is to use a sentence recognition test. When a reader reads a text, he/she comprehends by extracting the meaning of the sentences. In order to extract the meaning, the reader must decode the surface structure of the sentence (Sachs 1967; 1974). Graesser and Mandler (1975) developed a sentence recognition test that differentiated the participants' memory for a sentence's surface structure versus its meaning. The

surface structure of a text is the syntactic and phonological representation of a sentence. For example, consider the original sentence in the text as: "The restaurant is located on Young Ave., after the market." The surface structure of the sentence can be changed into: "The restaurant is located after the market, on Young Ave." The meaning of the original sentence can be changed without changing the surface structure. For example, the meaning of the original sentence can be changed into: "The restaurant is located on Young Ave., in front of the market." Both the surface structure and meaning of the original sentence can be changed, for example: "The restaurant is located in front of the market, on Young Ave." These changed sentences can be used as distractor items when assessing the reader's sentence recognition. Memory for surface structure is short lived (less than 1 minute) under many if not most reading experiences with expository texts (Sachs, 1967, 1974). However, recognition memory for surface structures increases when participants are instructed to pay attention to wording and phrasing of sentences during reading or when the discourse characteristics encourage the processing of surface features. Thus, the discrimination scores computed from the changed surface structures, changed meaning, and the original sentence in the sentence recognition test for surface structure and meaning can shed light on which of the two features (or both) a reader attends to and remembers while comprehending a text, especially with respect to participants' prior beliefs (Graesser & Mandler, 1975).

In addition to prior beliefs and general knowledge, texts that arouse a reader's affect may also increase the reader's memory for the text. Long (1989) investigated whether pragmatics and discourse style affected sentence recognition memory embedded in narratives. She found that sentence recognition memory increased as a function of the attitudinal statements of the characters in the story. In addition, when the sentences were emotionally expressive, the participants had substantial verbatim memories for them. This suggests that when a text is

emotionally charged, a reader's memory for that text may improve to not only remember what was said, but how the author of the text has said it.

Agent Roles in Agent Source Monitoring

In a classroom or a simulated intelligent tutoring environment with agents, the agents can act as sources of the disseminated information when they discuss the text that the learner is reading. The different roles the agents play may affect the reader's perceived validity of the disseminated information. For example, when learners were forced to pick a side between the disagreeing agents in D'Mello et al.'s 2015 study, they tended to side with the teacher agent rather than the peer agent, as would be expected if the reader considered the quality of the source (Goldman, et al., 2012). Conversely, research from online classrooms has also demonstrated that learners feel more at ease when they learn from a student discussion leader than from an instructor discussion leader (Rourke & Anderson, 2002). While information from an instructor is usually more reliable (Schermerhorn, Goldschmid, & Shore, 1976), instructors create an authoritarian presence which can actually impede learning (Bloxom, Caul, Fristoe, & Thomson, 1975; Kremer & McGuiness, 1988). This can occur when the discussion reverts to an initiate-respond-evaluate recitative structure (often involving the teacher asking shallow questions) or a traditional lecture, leaving the student passive and unreflective (Rourke & Anderson, 2002). The presence of a peer can counteract the tension of the instructor's authoritarian presence. Peer relations encourage a student to participate in idea sharing, competition, and provide a sense of belonging (Benard, 1993). Therefore, it should be noted that while the learner tends to regard the instructor as a more reliable source, the presence of the peer is needed to prompt deep learning. When a learner is given an emotionally charged text with controversial issues that challenge the learner's beliefs, the learner may be more motivated to perceive the information from the peer to

be more reliable if that information is consistent with the learner's beliefs. The relationship between perceived expertise, emotion, and prior beliefs is at the heart of the current investigation contributing to the source monitoring literature.

The Current Study

This current study combines research on cognitive disequilibrium, source monitoring, and memory of texts to investigate the effect of the different types of stimulated cognitive disequilibrium (i.e. clashes between prior belief and text, clashes between agent statements and text, clashes between prior beliefs and agent statements, and agents disagreeing with each other) on the memory and comprehension of texts with controversial topics. Participants read texts on controversial trendy current-events topics from mainstream news sites while listening to a concurrent conversation between a teacher agent and a student agent who discuss the text. Each text either argued for a pro position or a con position that either supported or did not support an issue (indicated as a "topic"). As the participants read the texts, they listened to the agents discussing the text and expressing points that may or may not disagree with each other and/or with the text. Each speech pair contained one statement spoken by the teacher agent and one statement spoken by the student agent on each page of each text. The spoken statements were either consistent or inconsistent with information in the text. The participants subsequently wrote summaries for the texts and took two multiple-choice tests. One of the multiple-choice tests assessed the participants' source memory, that is, which agent expressed a spoken comment (or neither). The other multiple-choice test assessed their sentence recognition memory for surface structure versus meaning. The main dependent measures of the current study are the summary scores, agent source memory scores, and recognition memory for sentence surface structure and meaning. Source memory scores were also used as a mediator of the summary scores and

sentence recognition memory scores to examine how and when participants attend to sources when confused. It is predicted that source memory should have at least a partial mediating effect on the memory measures. In addition, participants' affect states were accessed using an Affect Grid, which is a 9 x 9 questionnaire grid that simultaneously measures both valence and arousal (Russell, 2003) and were used as both a dependent variable and a covariate in separate models.

The current study primarily builds upon the designs on agents giving conflicting and/or incorrect claims about a text (D'Mello et al., 2015; Lehman et al., 2011). However, there are several differences with the current design. In the previous studies, the domain topic of "research methods" was expository and impersonal. The current study investigates whether similar patterns found from the previous research will apply to persuasive texts with topics that should elicit personal attitudes and affect. For example, in terms of the effect of agent roles, even though D'Mello et al.'s (2015) study found that the participants were biased towards agreeing with the instructor, the agent's source information likely mattered little to the participants because the participants' prior beliefs and world knowledge did not play a significant role in the stimulated confusion. This may be due to the participants' lack of intrinsic motivation to learn the emotionally flat topic of research methods and thus would not have an opinion on the topic. This is supported by D'Mello et al.'s (2015) finding that prior knowledge did not influence the participants' confusion, despite the participants' answering half (i.e. 51%) of the pretest questions correctly. This suggests that the participants did have some prior knowledge on the topic but the knowledge did not matter in stimulating cognitive disequilibrium. The current study utilizes texts with controversial topics to motivate the participants into becoming emotionally invested in learning the topics. The current design attempts to add the dimension of affect in

18

stimulating cognitive disequilibrium when the position of the text clashes with the participants'
prior beliefs.

The current design had four conditions similar to the study conducted by D'Mello et al.
(2015). However, because the current study used texts with controversial issues to consider the
factor of the reader's affect, the texts contained mostly the author's opinions and facts that are
still in dispute. Therefore, the current study considers the four agent speech conditions to be
either text-compatible (abbreviated as "+") or text-incompatible (abbreviated as "-") rather than
true or false. The agents' speeches consisted of argumentative opinions that were either
compatible or incompatible with the main ideas and supporting information presented in texts
rather than factual inferences. The T+S- condition consisted of the teacher agent speaking a
sentence statement that is compatible with the position of the text, and the student agent speaking
a sentence statement that is incompatible with the position of the text. The T-S+ condition
consisted of the teacher speaking a sentence statement that is incompatible with the position of
the text, and the student agent speaking a sentence statement that is compatible with the text. The
T+S+ condition consisted of both the teacher agent and the student agent each speaking a
statement that is compatible with the text. Finally, the T-S- condition consisted of both the
teacher agent and student agent each speaking a sentence statement that is incompatible with the
text.

In order to assess how the agents' opinionated disagreements affect text memory and text
comprehension, the current design presented the spoken disagreements after the learners had
read each page of the text, with the text still in view. This design constraint was exercised so that
the participants would be able to discriminate between the statements spoken by the agents and
the statements presented from the texts. The current study examined which agent source

participants remembered better when a pair of disagreeing spoken statements was presented. That is, it was important to consider whether the participants would continue to have a bias towards the teacher agent as a reliable source or overcome that bias when they listened to the position they agreed with in text with an emotionally charged topic.

The participants' agent source memory was tested after reading all the texts by using a source monitoring recognition test. The source monitoring recognition test asked the participants to choose the source (teacher vs. student vs. neither) that the presented speech statement was attributed to. The participants were presented the source memory exam after they had read all the texts to control for testing effects—that the participants began to attend to the agent speeches because they realized that they would be tested on them afterwards.

The participants' text memory was tested by a four alternative forced-choice test on sentence recognition memory that assesses their memory for surface structure and meaning (Graesser & Mandler, 1975). One of the four alternatives was an actual sentence from the text verbatim (S+ M+). The three other choices consisted the following: one item with the same surface structure but different in meaning (S+ M-), one item with different surface structure but the same meaning (S- M+), and the last item with both different surface structure and different in meaning (S- M-).

In addition to the memory measure on sentence recognition, the participants' deep comprehension and recall memory were assessed in their summary writing quality. They were asked to write a summary on what they understood regarding how the author of the text argued in terms of the position from the text. The participants were asked to write the summary after reading each text so that it was fresh in their minds.

In summary, the current study assesses the participants' sentence recognition memory for surface structure and meaning, their recall memory and comprehension for texts with their summary writing, and their memory for agent source for the particular agent statements. In addition, the study investigated the effect of prior beliefs on the controversial topics and the participants' affect after reading the text. The participants' prior beliefs were measured using a signal detection scale questionnaire. The participants' affect states were measured using an Affect Grid (Russell, 2003). Table 1 presents a full list of the measures of the current study.

The current study proposes two competing hypotheses regarding reading comprehension. According to the traditional argumentation research (Stein & Miller, 1990; Voss & Wiley, 2000), learners will attend to the text with the position that the learner agrees with on the controversial issue and better recall the statements from that text that agree with their position. From this perspective, the participants should score better on the sentence recognition exam for surface structure and meaning on the texts with the position they agree with, relative to the texts with the position they disagree with. Alternatively, according to the *Schema Pointer + Tag* theory (Graesser & Nakamura, 1982), the learners will put a "tag" in their memory representation for the statements with which they disagree. From this perspective, the participants should score better on the sentence recognition exam for surface structure and meaning on the texts with the position against the position of the participants, relative to the text with the position with which the participants agree.

The current study investigates both the effect of agent disagreements and the clashes of the position of the text with the participants' prior beliefs. There are two competing hypotheses involving the clashes that result in cognitive disequilibrium. One hypothesis is that the participants attend to sources either when the agents contradict each other through

disagreements, per D'Mello et al.'s (2015) research, or when the position of the text contradicts the participants' prior beliefs, per de Pereyra et al.'s (2014) research. There also may be an additive effect where the cognitive disequilibrium is greater when the participant experiences both clashes. An alternative hypothesis is that the participants only or predominately attend to sources when both clashes occur together: the text contradicts their prior belief and the agents disagree with each other.

Previous research reported that learners tend to trust the expertise of an instructor more than that of a peer (Lehman & Graesser, 2014; Rourke & Anderson, 2002). Thus, the learner may always be biased toward the teacher as a reliable source. Additionally, the saliency of remembering the argumentative statements and their source information should be higher for the teacher agent than the student agent, regardless of whether the teacher agent's position is for or against the participants' position on the controversial issue.

The other main effects involve participants' prior beliefs. According to the *Plausibility-Induced Source Focusing assumption* (de Pereyra et al., 2014), participants are more likely to attend to the source when the spoken information is in conflict with their world views. Based on this hypothesis, if the learners are emotionally invested in the position, the participants may remember the statements better when the statements agree with the participants' position regardless of the source. Previous research used general processing tree model (Riefer & Batchelder, 1995) to individually measure the participants' memory for recall, discrimination, and guessing. Their results suggested that the participants were not able to discriminate between the agent sources with accuracy (Feng, Cai, & Graesser, 2017). However, their materials were never on controversial topics, unlike the focus of the present study. In order to have a more

sensitive measure of the participants' metacognitive judgments on memory, the current study asked the participants to rate their confidence of the memory judgment on the agent sources.

Table 1. List of measures for the current study.

Construct	Measurements
Prior Beliefs	Prior Belief Questionnaire
Comprehension	Summary (50-100 words)
Comprehension	Sentence Recognition Test for Surface Structure and Meaning
Source Monitoring	Source Monitoring Recognition test
Emotional Response After Reading Text	Affect Grid

Method

Participants

A total of 128 participants randomly selected from the University of Memphis Psychology and Business Subject Pool participated in this study. A power analysis was conducted to determine the number of participants. For each marginal comparison of a two-level factor, and assuming the Cohen's effect size of half standard deviation, the estimated sample size would be $16/d^2 = 64$ per group, which resulted in total 128 participants for this study. Furthermore, since we had 32 combinations of various factors, with 128 participants, we would have 4 participants to ensure robust estimate of effect in each cell. All participants were 18-years-old or older ($M = 21.93$, $SD = 5.86$). Approximately 55.5% of the participants identified themselves as female and 44.5% of the participants identified themselves as male. Approximately 42.2% of the participants identified themselves as African/African American, 39.8% identified as White/European Descent, 8.6% identified as Asian, 4.7% identified as Hispanic/South American and 4.7% identified as Middle Eastern. In addition, the largest

percentage of the participants identified themselves as "independent/non-politically affiliated" (26.6%), the second largest were the "moderate/centrist" (17.2%), the third identified as "democrat" (14.1%), fourth identified as liberal/classical liberal (10.9%), fifth identified as "socialist/feminist/green" (9.4%), sixth identified as "libertarian" and "republican" (7.8% for both), and the fewest group identified as "very conservative/traditionalist" (6.3%).

Before the experiment, all participants were given a written informed consent in compliance with a protocol approved by the Institutional Review Board of the University of Memphis. The consent form included purpose and details regarding the experiment, as well as any risks and/or benefits for participating. All participants were given a debriefing informing them about the manipulations of the study as well as reiterating the goal of the study once they completed the experiment.

Experiment Design

The experiment used a 2 text position (text pro or text con) x 4 agent-to-text compatibility (T+S-, T-S+, T+S+, and T-S-) within-subject design. Participants' prior beliefs regarding the topics were evaluated based on what the participants reported on a questionnaire before the interaction starts. The questionnaire asked the participants: "How strongly do you support X?" For example: "How strongly do you support the Trump Presidency?" The participants were asked to choose a side (pro or con) on 1 to 6 anchored scale:

6. Strongly support   5. Moderately support   4. Somewhat support

3. Somewhat against   2. Moderately against   1. Strongly against

All participants interacted with the learning environment AutoTutor (Cai, Graesser, & Hu, 2015) with two conversational talking head agents. One agent assumed the role of the teacher, and the other assumed the role of the student. The agents had a single-sentence per turn conversation back and forth while the participant read the text. The participants' reading time per page in seconds was recorded. On average, the participants took 66 minutes ($SD = 18$) to complete the experiment.

After the participants completed the *AutoTutor* phase for each text, they were tested on their comprehension of the texts through summarization. After the participants completed the AutoTutor phase for all four texts, they were tested on their sentence recognition and source memory. In addition, they completed a brief demographics and social desirability questionnaire before the start of the experiment. The demographics asked the participants their age, gender, race/ethnicity, and political affiliation. The social desirability scale was a short 13-question version of the Marlowe-Crowne questionnaire (Ballard, 1992).

Finally, to assess the participants' emotional state after they read each text, the participants were presented with an affect grid to measure their valence and arousal. The affect grid is a 2-dimensional single item affect measurement questionnaire consisting of a 9×9 (valence $\times$ arousal) grid, as shown in Figure 2. The affect grid has a correlation of .89 or higher with other scales of affect (Russell, 2003). The benefit of the affect grid as opposed to a conventional questionnaire is that participants are minimally influenced by emotion labels and they can simultaneously provide an indicator of both valence (on the horizontal plane) and arousal (on the vertical plane), which are two universal dimensions of basic emotions according to Russell (2003). The participants were asked to select the cell that best represent their current

affective state. The measure served as a manipulation check to assess whether the topics of the current design resulted in high emotional arousal from the participants.
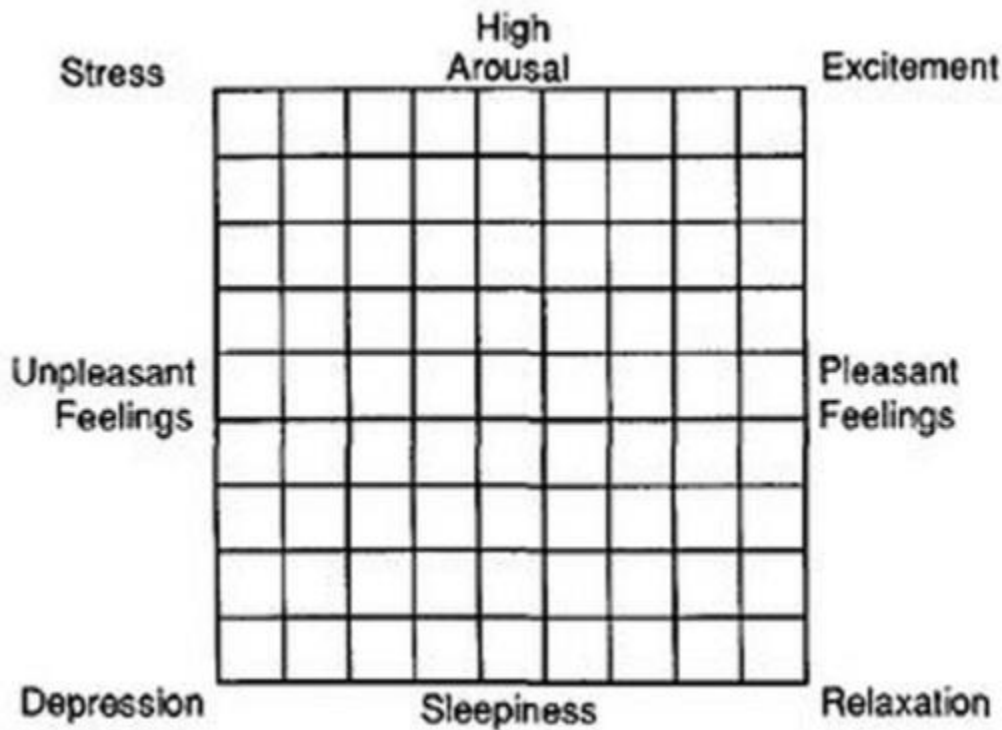

Figure 2. Affect Grid (Russell, 2003).

Materials

The *AutoTutor* interface with two agents (Graesser, Li, & Forsyth, 2014) was used in the learning environment with which the participants interacted. The learning environment had an animated African-American female agent named Cristina, and a racially ambiguous/Latino male agent named Jordan. The agent roles (teacher or student) were counterbalanced. That is, half of the participants received Cristina and the other half of the participants received Jordan as the teacher. The interface that was used for the current experiment is shown in Figure 3. An example text and statements of the agents is also presented in Figure 3.

There's no question that Donald Trump is a proud American. In fact, there are only two things Trump consistently promotes—himself and the country he lives in. However, Donald Trump has shown that he has no interest in the wider world and what goes on beyond America's borders. And while that might be fine when you're a billionaire businessman, it could be disastrous if he becomes president. After all, much of the president's role involves foreign policy. From brokering peace in the Middle East to providing aid during humanitarian crises and asserting military power where needed, being president requires an understanding of foreign relations, politics and diplomacy. And Donald Trump has none of those things. Interestingly, the last president to be accused of being a disaster on foreign relations was George W. Bush.

Donald Trump has never been an elected official. He has never held elected office of any kind. He's never had to broker political compromise, give political speeches or pour over political briefing notes. So why on earth, with no experience in politics, would people think he'd make a good president? Sure, Donald Trump is a shameless self-promoter and experienced businessman. But Washington, D.C., is not Wall Street, and getting things done in the nation's capital takes more than a big personality and strong will. If Donald Trump is elected president, he'll quickly have to learn how to deal with his Democratic counterparts in the House of Representatives and Senate. He'll also have to get used to the pace and minutiae of government—where things can move extremely slow. And he'll have to get used to giving speeches on topics he may not be interested in—like housing for the poor, education and farm subsidies. Being in office, after all, is not the same as campaigning.

Figure 3. Example Screenshots of Learning Environment in the Four Agent-to-Text Compatibility Conditions.

A total of 4 controversial topics and 8 texts were presented using AutoTutor. For each topic, there were two texts with different positions (supporting the topic or against the topic). The topics of the texts covered were *Black Lives Matter Movement*, *Men's Rights Movement*, *Trump Presidency*, and *Syrian Refugee Crisis*. Appendix A shows an example text and interaction. These topics were selected from the Pew Research Center on the top voting issues of 2016 (http://www.people-press.org/2016/07/07/4-top-voting-issues-in-2016-election/). The texts were selected from reputable, public, main stream sources such as *The Washington Post* and *Time* magazine.

Measures

Summary Scores. The participants were instructed to write a short 50 to100-word summary regarding the content arguments and the position of the issues presented in the text. They were asked to write each text summary after reading each text and listening to the agents' talk. The summary measured both the participants' comprehension of the material and their recall memory. The quality of summaries was determined as a single holistic score on a 1 to 4 continuous scale. The single score reflects how the participant understood and recalled accurately the main ideas and supporting arguments for the main ideas. The summary qualities of all 128 participants were scored by two expert graders. Lower scores indicated low comprehension. There was a high inter-rater reliability of .81, assessed by Cronbach's Alpha.

Source monitoring recognition test. Participants were given a source monitoring recognition test at the end of the AutoTutor interaction. The test presented the participants with a total of 64 speech statements from the agents. The participants were asked to choose from options: Teacher said it, Student said it, or Neither said it. Half of the statements (32 items) were

statements that the participants have heard from the agents. The other half of the presented

statements (32 items) were presented as new items. Furthermore, half of the participants heard

the first 32 items as old items, and the other half of the participants heard the second 32 items as

old items. For the agent speech conditions of T+S- or T-S+, the new items were the old items

that the other half of the participants received in these two speech conditions (and vice versa).

Similarly, for the agent speech condition of T+S+, the new items were the items from the old

items from the condition T-S- that the other half of the participants received. And finally, for the

agent speech condition T-S-, the new items were the old items from the condition T+S+ that the

other half of the participants received (See Appendix B for an example). Participants were also

asked to rate their level of confidence by selecting from the following options on a 1 to 6

anchored scale:

1. Not confident with my judgment at all.

2. Not very confident with my judgment.

3. A little unconfident with my judgment.

4. A little confident with my judgment.

5. Moderately confident with my judgment.

6. Completely confident with my judgment.

Recognition Memory Test for Surface Structure and Meaning of Sentences. Participants

were given a sentence recognition exam that consisted of four alternative forced-choice

questions. Each question asked the participant to identify the verbatim sentence that appeared in

the text that they read (i.e. same surface structure and meaning; S+M+). The other three

alternatives had one sentence that consisted of the same surface structure but with the meaning of the sentence altered (S+M-), one sentence that altered the surface structure (such as a word, word order or syntax) but had the same meaning (S-M+), and one sentence which had both different surface structure and an altered meaning (S-M-). Appendix C shows example test items for one text. There were 4 questions per text. Each participant received a total of 16 questions for the four texts that they read. Each question was tailored to a specific page of the text and was presented in an order corresponding to the page of the text that was presented.

Manipulations

Text Positions. There were a total of 8 texts. There were 2 texts for each topic. One of the two texts supported the topic, and the other text did not support the topic. The lengths of the texts were between 699 – 1111 words (*M* = 931.38, *SD* = 154.40). Each text pair (i.e. a pro text and a con text, both on the same topic) were selected to match in their lengths. Each participant read a total of four texts, one text per topic, and one on each topic. Two of the texts they received supported the topic, and the other two texts they received did not support the topic. The presentation of the order of the texts was counterbalanced using a Greco-Latin Square (see Appendix D for the full counterbalancing table).

Agent-to-Text Compatibility. During the reading section of the experiment, both agents spoke single-sentence statements. In one text, both agents spoke in agreement with the position of the text (T+S+). In another text, the teacher agent spoke in agreement with the position of the text and the student agent spoke in disagreement with the position of the text (T+S-). In the third text, the teacher agent spoke in disagreement with the position of the text and the student agent spoke in agreement with the position of the text (T-S+). In the fourth text a participant received

31

both agents speaking in disagreement with the position of the text (T-S-). Consequently, two of the texts that any one student read were in support of the topic and two of the texts they read did not support the topic. One pair of argumentative statements were presented on each page. Since there were 4 pages per text, a total of 8 spoken statements per text were presented to the participant. Appendix A presents an example text and the content of the agents' spoken statements.

Counterbalancing. The experimental design involved many combinations of variables, so constructing an adequate counterbalancing scheme to insure the integrity of the results was paramount. The following list is a summary of the counterbalancing scheme of the current design, whereas Appendix D shows the full counterbalancing table:

1. Presentation of topic will rotate in accordance with a Greco-Latin Square.
2. Half of the participants will have Cristina as the teacher agent and Jordan as the student agent. The other half of the participants will have Jordan as the teacher agent and Cristina as the student agent.
3. The order of who speaks first in the agent speech manipulation will be counterbalanced. Specifically, the agents will take turns who speaks first within each text as the participants read each page within each text.
4. The order of the presentation of the four agent speech conditions will be counterbalanced.

Procedure

Before the experiment started, the participants were given a consent form for agreeing to participate in the study. The participants were first instructed to complete the demographics and

the social desirability questionnaire. They subsequently completed the prior belief questionnaire for all 4 topics. Their answers were recorded by the computer using the survey service *Qualtrics*, an online software for conducting surveys and recording survey data. After that, the participants were directed to AutoTutor to start the interaction with the animated agents. They were first instructed by the agents regarding their task. The agents informed the participants that they are to read each of the four texts in a page-by-page format. They were told that they would read them to assess their reading abilities. In addition, they were told to pay attention to the discussion between the agents. The participants were told to click on the "Next Page" button after they finish reading each page. This prompted the agents to speak their agreements or disagreements in their respective conditions. Participants were instructed not to click the "Next Page" button before they have finished reading that page. After reading the entire text, participants were instructed to complete the Affect Grid regarding their current affect state and write a summary of what they had just read. Figure 4 shows the overall order of events for the experiment. The purple area indicates repetition when the participant moves to the next text.
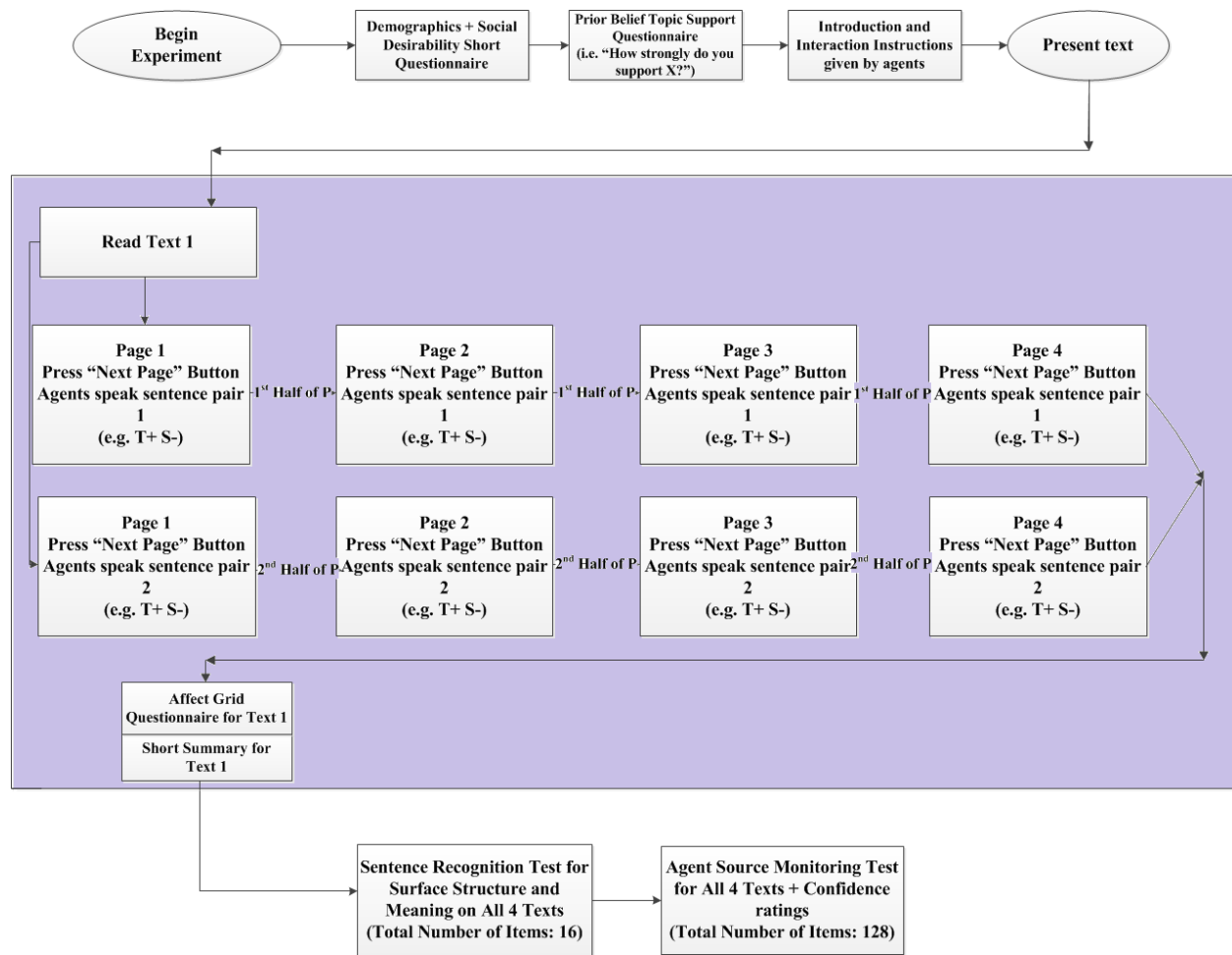
Figure 4. Order of Events for the Procedure of the Experimental Design.

After the participants finished reading all four texts, they completed the sentence recognition test on sentences explicitly presented in the texts they had read. The subsequent step was completing the source monitoring test. Each of the tests presented questions in the order of the pages were presented to the participants. For the source monitoring test, each page had 4 speech statements, with the four statements randomized per page for each participant separately (but the order of each page stayed the same as what had been presented). There were a total of 16 questions in the four alternative forced-choice sentence recognition test for surface structure and meaning.

There were a total of 64 questions asking the participant to identify the sources of statements in the source monitoring test (Teacher, Student, or Neither). In addition, each question in the source monitoring test requested the participants to rate their confidence in selecting the correct source on the 6-point scale presented earlier.

Results

The design involved a repeated-measures and a nested structure of data (conditions within text, texts within individuals) so a mixed-effects modeling approach was adopted for all analyses. Mixed-effects modeling has been recommended for this type of within-subject data sets (Goldstein, 2011) with complex counterbalancing, nesting, and variability across both materials and participants. The models would account for the random effects that can influence the fixed effects on dependent variables. Logistic or continuous analysis of variance mixed effect models were conducted depending on whether the dependent variables were binary versus continuous. An alpha level of .05 was adopted for all analyses. As the analyses used mixed-effect models, the degrees of freedom used the Kenward-Roger degrees of freedom approximation (Kenward & Roger, 1997) which adjusts for bias for each model. The random effect for all data sets was participants (128 levels). The independent (fixed effects) variables for all analyses were: prior beliefs on the position of topic (sometimes abbreviated as topic support), compatibility of text position with prior beliefs (compatible versus incompatible based on participants' self-report), agent-to-text compatibility (T+S-, T-S+, T+S+, T-S-), compatibility of the participants' prior belief to the position of the teacher, compatibility of the participants' prior belief to the position of the student, and whether agents agreed or disagreed with each other. Agent agreeing included T+S+ and T-S-, whereas agent disagreeing included T+S- and T-S+.

The agent agreement/disagreement variable was simply a binary grouping of the 4-agent-to-text compatibility variable, so separate models must be conducted when using each. The independent variables of prior beliefs on the position of the topic and of compatibility of text position with prior beliefs were both binary conditions. They were calculated based on the ratings of the prior beliefs questionnaire that that participants filled out before reading the texts. The questionnaire was an anchored 1-6 scale so the participants who self-reported scores of 1, 2, or 3 were grouped into the condition of "not supporting topic," whereas the participants who self-reported scores of 4, 5, or 6 were grouped into the condition of "topic support." The binary conditions of topic support and compatibility of text position to prior beliefs were also computed based on this metric, with the addition of taking the position of text into account. The participants who self-reported scores of 1, 2, or 3 were grouped into the condition of "prior belief incompatible to text position" whereas the participants who self-reported scores of 4, 5, or 6 were grouped into the condition of "prior belief compatible to text position." The frequency data revealed that the participants tended to be heavily biased in reporting support for the topic than not support (support: $N = 362$; not support: $N = 150$). This was especially salient for any topic that the participant did not have a strong opinion on, i.e. the scores of 3 or 4 ($N = 186$ for a score of 4 and $N = 50$ for a score of 3).

In the analyses of source memory data, logistic mixed effect models were conducted using source memory scores (1 or 0) as the dependent variable. For the recognition memory test, mixed effect analysis of variance models were conducted when the dependent variables were: participants' discrimination proportion scores for surface structure (continuous variable) and participants' discrimination proportion scores for meaning (continuous variable). In addition, participants' summary scores and participants' reading time were also continuous dependent

variables. Finally, valence and arousal (based on affect grid scores) were used both as continuous dependent variables and as covariates for manipulation check. Race, gender, political affiliation, and social desirability were also tested as covariates in all models to rule out potential confounds. It has been found for all models that the models with these covariates were not significantly better models than the models without these covariates. This suggests that valence and arousal, race, gender, political affiliation, and social desirability did not significantly affect the fixed factors in predicting the dependent variables.

Topics Support Ratings. On average, the participants rated themselves to be significantly less supportive of the topic regarding Trump Presidency than the other three topics ($F(3, 508) = 94.08$, $p < .001$; $MSe = 153.74$, $\eta^2 = .357$). The means were $2.27(SD = 1.54)$, $4.66(SD = 1.31)$, $4.21(SD = 1.10)$, and $4.41(SD = 1.10)$ for the Trump Presidency, Black Lives Matter, the Men's Rights Movement, and the Syrian Refugee Crisis, respectively. The latter three topics did not significantly differ.

Affect. On average, the participants rated the texts to be moderately pleasant ($M = 5.26$, $SD = 2.29$), and above average in terms of arousing ($M = 4.58$, $SD = 2.15$). However, neither the ratings on valence nor arousal significantly differed from each other. Participants rated the texts on the Trump Presidency to be the highest on pleasantness ($M = 5.35$, $SD = 2.29$) and the texts on the Black Lives Matter Movement as the lowest on pleasantness ($M = 5.03$, $SD = 2.26$). The participants found the texts on the Men's Rights Movement to be the most exciting ($M = 4.79$, $SD = 2.13$) and the texts on the Syrian Refugee Crisis to be the least exciting ($M = 4.45$, $SD = 1.88$).

For affect valence ratings, there was a significant main effect of compatibility of participants' prior beliefs to the position of the text ($F(1, 430.82) = 7.24$. $p = .007$, $MSe = 21.21$, $\eta^2 = .010$). Participants who reported their prior beliefs compatible with the position of the text had higher valence ratings ($M = 5.51$, $SD = 2.25$, $N = 256$) than participants who reported their prior beliefs incompatible with the position of the text ($M = 5.00$, $SD = 2.31$, $N = 256$). In addition, there was also a significant main effect of whether agents agreed or disagreed with each other on valence ratings ($F(1, 390.10) = 4.87$, $p = .03$, $MSe = 14.26$, $\eta^2 = 0.006$). When participants listened to the agents agreeing with each other, participants had significantly higher valence ratings ($M = 5.43$, $SD = 2.34$, $N = 256$) than the participants who listened to the agents disagreeing with each other ($M = 5.09$, $SD = 2.23$, $N = 256$). No significant interactions between the prior belief conditions and the agent speeches conditions were found.

For affect arousal ratings, there was a significant main effect of compatibility of participants' prior beliefs to the position of the text ($F(1, 455.70) = 6.02$. $p = .01$, $MSe = 19.91$, $\eta^2 = .009$). Participants who reported their prior beliefs compatible with the position of the text had higher arousal ratings ($M = 4.78$, $SD = 2.17$, $N = 256$) than participants who reported their prior beliefs incompatible with the position of the text ($M = 4.38$, $SD = 2.11$, $N = 256$). There was no significant effect of agent speeches or significant interactions.

Discrimination Memory for Surface Structure. Memory discrimination scores for surface structure were computed based on Graesser and Mandler (1974). The probability of remembering surface structure was computed as the proportion of choosing (S+M-) + (S+M+) over the total number of questions per text. There was a total of 4 questions per text, so the proportion scores for surface structure would be either 0, 0.5, 0.75, or 1. The probability of .50 as chance level was used in all computations.

On average, participants correctly identified 0.66 ($SD = 0.25$) of the surface structure from the original sentences of the text they read, which was significantly above chance level ($t(511) = 14.30$, $p < .001$). Summary and descriptive statistics for participants' sentence recognition for surface structure are presented in Table 2.

There was a significant main effect of prior beliefs on the position of topic on identifying the original surface structure when controlling for agent-to-text compatibility and compatibility of text position with prior beliefs ($F(1, 463.57) = 4.05$. $p = .04$, $MSe = 0.23$, $\eta^2 = .007$). Participants who reported that they did not support the topic identified a higher portion of the original surface structure of the sentences of the text they read ($M = 0.69$, $SD = 0.26$, $N = 150$) than the participants who reported that they supported the topic ($M = 0.65$, $SD = 0.25$, $N = 362$). There was also a significant main effect of compatibility of participants' prior beliefs to the position of the teacher ($F(1, 245.47) = 5.94$. $p = .02$, $MSe = 0.27$, $\eta^2 = .002$). Participants who reported their prior belief incompatible with the position of the teacher had higher recognition memory for surface structure of the sentences ($M = 0.68$, $SD = 0.25$, $N = 255$) than participants who reported their prior belief compatible with the position of the teacher ($M = 0.64$, $SD = 0.26$, $N = 257$). In addition, when controlling for prior beliefs on the position of topic and compatibility of text position with prior beliefs, there was a significant main effect of whether agents agreed or disagreed with each other on identifying the original surface structure ($F(1, 380) = 4.09$, $p = .04$, $MSe = 0.24$, $\eta^2 = 0.007$). When participants listened to the agents disagreeing with each other, participants identified a higher portion of the original surface structure of the sentences of the text they read ($M = 0.68$, $SD = 0.25$, $N = 256$) than the participants who listened to the agents agreeing with each other ($M = 0.64$, $SD = 0.26$, $N = 256$). No significant interactions between the prior belief conditions and the agent speeches conditions

were found. These results confirm the prediction that cognitive disequilibrium can lead to better memory for the surface structure in presented information.

Table 2. Descriptive statistics for discrimination memory for surface structure of sentences.

Prior Belief Conditions		Agent Speech Conditions		
		Agents agreeing	Agents disagreeing	N Total
Prior belief supporting the topic	Prior belief compatible with text	0.57(0.27)	0.68(0.25)	181
	Prior belief incompatible with text	0.67(0.25)	0.68(0.23)	181
Prior belief not supporting the topic	Prior belief compatible with text	0.71(0.24)	0.68(0.28)	75
	Prior belief incompatible with text	0.67(0.27)	0.70(0.25)	75

Descriptive statistics were examined on the relationship between three interactive sources of cognitive clashes: prior belief incongruent with the position of the text, agent disagreement, and agent's position incongruent with the position of the text. The participants had the highest recognition memory for surface structure when they experienced either of the following interacting clashes: when their prior beliefs were incongruent with the position of the text, when the agents disagreed with each other, but teacher's position was compatible with the text ($M =$ 0.70, $SD = .23$), or when their prior beliefs were incongruent with the text, the agents agreed with each other, but the teacher's position was incompatible with the text ($M = 0.70$, $SD = 0.23$). On the other hand, the participants' lowest recognition for surface structure occurred when their prior beliefs were congruent with the text and the agents agreed with each other, regardless of the teacher's position (teacher compatible: $M = 0.61$, $SD = 0.27$; teacher incompatible: $M = 0.62$, $SD = 0.27$).

Perhaps the lack of effect from compatibility of text position with prior beliefs was due to the inclusion of the biased topic support ratings. To investigate this further, two sets of separate analyses were conducted. The first set investigated participants' recognition memory on sentences' surface structure when the participants reported strong opinions of the topic (prior belief ratings of 1, 2, 5, and 6). The second set investigated participants' recognition memory on sentences' surface structure when participants rated more neutral opinions of the topic (prior belief ratings of 3 and 4).

When using only the data with extreme prior belief ratings (i.e. ratings of 1 and 2 for not support, ratings of 5 and 6 for support), there was a significant main effect of compatibility of text position with prior beliefs on identifying the original surface structure ($F(1, 209.23) = 6.47$, $p = .01$, $MSe = 0.29$, $\eta^2 = 0.01$). When participants' prior belief about the position of the topic was strongly incongruent with the position of the text, they had a significantly higher recognition of the text's surface structure of the sentences they read ($M = 0.71$, $SD = 0.23$, $N = 136$) than if their prior belief about the topic was strongly congruent with the position of the text ($M = 0.66$, $SD = 0.26$, $N = 140$). There was also a significant main effect of prior beliefs on the position of topic on identifying the original surface structure ($F(1, 208.86) = 7.40$, $p = .007$, $MSe = 0.33$, $\eta^2 = 0.02$). When participants' prior beliefs were strongly not supportive of the topic, they had a significantly higher recognition of the text's surface structure of the sentences they read ($M = 0.73$, $SD = 0.23$, $N = 100$) than if their prior belief was strongly supportive of the topic ($M = 0.66$, $SD = 0.25$, $N = 176$). However, there was no significant main effect of whether the agents agreed with each other or the agent speech in these conditions. There was a marginal significant interaction between whether agents agreed with each other and participants' prior beliefs on the position of topic ($F(1, 269.03) = 3.80$, $p = .052$, $MSe = 0.17$, $\eta^2 = 0.004$). However, a post hoc

test suggested that the only significant pair was when participants' reported not supportive of the topic ($M = 0.74$, $SD = 0.22$, $N = 55$) versus when they reported being supportive ($M = 0.63$, $SD = 0.25$, $N = 86$) while listing to both agents agreeing with each other ($t(258.86) = 3.32$, $p = .006$). Therefore, when participants reported having strong or solid opinions of the topic, their prior beliefs, rather than agents, significantly affected their surface structure memory of the text they read.

Consider next the data with more neutral prior belief ratings (i.e. ratings of 3 for not support, ratings of 4 for support). In this case, there was no significant main effect of either compatibility of text position with prior beliefs or of their prior beliefs on the position of topic on the participants' identifying the original surface structure. However, there was a significant main effect of whether agents agreed or disagreed with each other on the participants' identifying the original surface structure ($F(1, 184.94) = 3.87$, $p = .05$, $MSe = 0.26$, $\eta^2 = 0.02$). When participants encountered agents disagreeing with each other over the position of the text, they had a significantly higher recognition of the text's surface structure of the sentences they read ($M = 0.67$, $SD = 0.25$, $N = 121$) than if they had encountered agents agreeing with each other over the position of the text ($M = 0.60$, $SD = 0.28$, $N = 115$). There were no significant interactions between the agent conditions and the participants' prior beliefs. These results suggested that when participants' opinion on the topic was more neutral, listening to the agents debating about the position of the topic from the text had a significant impact over the participants' ability to identify the original surface structure of the texts they read. As the participants' prior beliefs were more neutral regarding the topic, their prior belief did not significantly affect their ability to identify the original surface structure of the text they read.

We also explored the effect of topic on the participants sentence recognition. When topic was added into the model as a fixed factor, there was a significant main effect of topic on the participants' sentence recognition memory for the surface structure of the text ($F(3, 391.49) = 4.02$, $p = .008$, $MSe = 0.23$, $\eta^2 = 0.006$). A post hoc test revealed that the topic regarding the Trump Presidency (of which topic they were least supportive of) had significantly higher recognition memory for surface structure ($M = 0.73$, $SD = 0.22$, $N = 128$) than either the texts with the topic of Black Lives Matter Movement ($M = 0.65$, $SD = 0.27$, $N = 128$; $t(410.19) = 2.62$, $p = .04$), Men's Rights Movement ($M = 0.64$, $SD = 0.25$, $N = 128$; $t(408.60) = 2.94$, $p = .02$), or the Syrian Refugee Crisis ($M = 0.63$, $SD = 0.26$, $N = 128$; $t(410.19) = 3.24$, $p = .007$). The sentence recognition memory for surface structure on the latter three topics did not significantly differ from each other.

As a follow-up, four separate models were conducted for each of the topics. Surprisingly, only when participants read texts regarding the Syrian Refugee Crisis, there was a significant main effect of prior beliefs on the position of topic on the participants' identifying the original surface structure ($F(1, 122) = 4.01$, $p = .05$, $MSe = 0.27$, $\eta^2 = 0.030$). Contrary to previous findings above, when participants reported support for the topic of the Syrian refugees, they had a significantly higher sentence recognition memory for surface structure ($M = 0.64$, $SD = 0.26$, $N = 110$) than when they reported not supportive of the topic of the Syrian refugees ($M = 0.51$, $SD = 0.26$, $N = 18$). There was a significant main effect of whether the agents agreed or disagreed with each other on the participants' identifying the original surface structure ($F(1, 126) = 6.11$, $p = .01$). While reading texts on the topic of the Syrian Refugee Crisis, when participants encountered the agents disagreeing with each other over the text, they had significantly higher sentence recognition memory for surface structure ($M = 0.68$, $SD = 0.25$, $N = 64$) than if the

participants encountered the agents agreeing with each other over the text ($M = 0.57$, $SD = 0.27$, $N = 64$). There was also a significant main effect of the agent speech conditions on the participants' identifying the original surface structure ($F(3, 116) = 3.15$, $p = .03$, $MSe = 0.21$, $\eta^2 = 0.07$). While reading texts on the topic of the Syrian Refugee Crisis, participants had significantly higher sentence recognition memory for surface structure when they listened to the teacher against the position of the text while the student for the position of the text ($M = 0.74$, $SD = 0.23$, $N = 32$) than the two conditions where the agents agreed with each other (T-S-: $M = 0.58$, $SD = 0.29$, $N = 32$, $t(124) = 2.56$, $p = .05$; T+S+: $M = 0.56$, $SD = 0.25$, $N = 32$, $t(124) = 2.79$, $p = .03$). There was no significant effect between the two conditions where the agents agreed with each other or between the two conditions where the agents disagreed with each other. There was also no interaction between the manipulation of the agents and participants' self-reported prior belief groups.

Discrimination Memory for the Meaning of Sentences. Memory discrimination scores for meaning were computed based on Graesser and Mandler (1974). The probability of remembering meaning was computed as the proportion of choosing (S-M+) + (S+M+) over the total number of questions per text. The probability of .50 as chance level was used in computations that assessed whether the recognition memory differed from chance.

On average, participants correctly identified the meaning 0.74% ($SD = .25$) of the original sentences of the text they read, which was significantly above chance level ($t(511) = 21.53$, $p < .001$).). Summary and descriptive statistics for participants' sentence recognition for meaning are presented in Table 3.

There was a significant main effect of compatibility of text position with prior beliefs on identifying the original meaning of the text ($F(1, 447) = 5.82$. $p = .02$, $MSe = 0.26$, $\eta^2 = .004$). Participants who reported prior belief incompatible with the text ($M = 0.75$, $SD = 0.23$, $N = 256$) had significant higher recognition memory for meaning than the participants who reported their prior belief compatible with the text ($M = 0.72$, $SD = 0.27$, $N = 256$). There was no significant effect of participants' prior belief on the position of the topic nor were there effects of the agent speech manipulation conditions. In addition, there were no significant interactions between the prior belief conditions and the agent speech conditions. The results did not significantly differ when separate models were constructed either using only the data with extreme prior belief ratings (i.e. ratings of 1 and 2 for not support, ratings of 5 and 6 for support) or using the data with more neutral prior belief ratings (i.e. ratings of 3 for not supporting, ratings of 4 for support).

Table 3. Descriptive statistics for discrimination memory for meaning of sentences.

Prior Belief Conditions		Agent Speech Conditions		
		Agents agreeing	Agents disagreeing	N Total
Prior belief supporting the topic	Prior belief compatible with text	0.71(.26)	0.67(0.27)	181
	Prior belief incompatible with text	0.78(.21)	0.76(0.23)	181
Prior belief not supporting the topic	Prior belief compatible with text	0.78(0.28)	0.81(0.22)	75
	Prior belief incompatible with text	0.67(0.24)	0.73(0.23)	75

Descriptive statistics were examined on the relationship between three interactive sources of cognitive clashes: prior belief incongruent with the position of the text, agent disagreement,

and agent's position incongruent with the position of the text. Participants had the highest recognition memory for meaning when they experienced either the following interacting clashes: their prior beliefs were incongruent with the position of the text, the agents disagreed with each other, and teacher's position was incompatible with the text ($M = 0.77$, $SD = 0.22$). On the other hand, the participants' lowest recognition for meaning occurred when their prior beliefs were congruent with the text and the agents disagreed with each other, but the teacher's position was compatible with the text ($M = 0.69$, $SD = 0.28$).

When topic was added into the model as a fixed factor, there was a significant main effect of topic on the participants' sentence recognition memory for meaning of the text ($F(3, 381) = 4.40$, $p = .02$, $MSe = 0.19$, $\eta^2 = 0.02$). A post hoc test revealed that the topic regarding the Black Lives Matter Movement had significantly lower recognition memory for meaning ($M = 0.68$, $SD = 0.26$, $N = 128$) than either the texts with the topic of the Men's Rights Movement ($M = 0.75$, $SD = 0.23$, $N = 128$; $t(381) = -2.69$, $p = .04$), Trump Presidency ($M = 0.76$, $SD = 0.25$, $N = 128$; $t(381) = -2.99$, $p = .02$), or the Syrian Refugee Crisis ($M = 0.76$, $SD = 0.24$, $N = 128$; $t(381) = -3.14$, $p = .01$). The sentence recognition memory for meaning of the latter three topics did not significantly differ from each other.

As a follow-up, four separate models were conducted for each of the topics. Contrary to the analysis conducted on participants' recognition for surface structure, the results show that only when the participants read texts regarding the Black Lives Matter Movement was there a significant main effect of compatibility of text position with prior beliefs on the participants' recognition memory for the original meaning ($F(1, 126) = 12.03$, $p < .001$, $MSe = 0.74$, $\eta^2 = 0.09$). Consistent with the previous analysis above, when participants' prior beliefs were reported to be incompatible with the position of the texts on the topic of the Black Lives Matter

Movement, they had a significantly higher sentence recognition memory for meaning ($M = 0.75$, $SD = 0.23$, $N = 64$) than when their prior belief was compatible with the position of the texts on the topic of the Black Lives Matter Movement ($M = 0.60$, $SD = 0.27$, $N = 64$). There was a significant interaction between whether the agents agreed with each other and participants' prior beliefs on the position of the topic affecting participants' recognition memory for the original meaning on the topic of the Trump Presidency ($F(1, 122) = 5.04$, $p = .03$). However, Tukey Post Hoc test showed no significance in any of the pairwise tests. There was also no main effect of either the agent agreeing/disagreeing conditions or participants' prior beliefs on the position of the topic when each were tested in separate individual models.

Summary scores. The quality of the participants' summaries was scored by two independent expert raters on whether the participants have correctly understood the arguments for the positions that were presented in each of the texts. The scores were on a 1 to 4 continuous scale, with lower scores indicating low comprehension. A Cronbach's Alpha for inter-rater reliability was conducted. The result showed a high inter-rater reliability of .81 on a total of 512 summaries between the two raters. The two independent scores were then averaged for each participant's summary score.

The participants' mean summary score was 2.14 ($SD = .85$). The distribution of the summary scores suggested that most participants received a summary score of 1 and 2. Using the third central moment, skewness is .29 with a kurtosis of -.89. Summary and descriptive statistics for participants' summary scores are presented in Table 4.

There were no statistically significant main effects of independent variables prior beliefs on the position of topic, compatibility of text position with prior beliefs, agent-to-text

compatibility, and whether agents agreed or disagreed with each other on participants' summary scores when each independent variable was conducted as separate models. There was, however, a significant interaction between compatibility of text position with prior beliefs and whether agents agreed with each other on the summary scores ($F(1, 463.89) = 7.44. p = .007$, $MSe = 3.65$, $\eta^2 = .02$). When the participants reported their prior beliefs as congruent with the position of the text, the participants who listened to the agents disagreeing with each other ($M = 2.36$, $SD = .89$, $N = 125$) had significant higher summary scores than the participants who listened to the agents agreeing with each other ($M = 2.02$, $SD = .83$, $N = 125$; $t(428.78) = 3.22$, $p = .008$). There was also a significant interaction when analyzing agreement among agents and the compatibility of participants' prior beliefs to the teacher agent's position ($F(1, 428.07) = 10.78. p = .001$, $MSe = 5.21$, $\eta^2 = .02$). When participants reported their prior belief as compatible with the position of the teacher agent, the participants who listened to the agents agreeing with each other ($M = 2.36$, $SD = .85$, $N = 117$) had significantly higher summary scores than the participants who listened to the agents disagreeing with each other ($M = 1.98$, $SD = .78$, $N = 138$; $t(404.18) = 3.64$, $p = . 002$).

Table 4. Descriptive statistics for summary scores.

Prior Belief Conditions		Agent Speech Conditions		
		Agents agreeing	Agents disagreeing	N Total
Prior belief supporting the topic	Prior belief compatible with text	2.31(.86)	1.97(.75)	181
	Prior belief incompatible with text	2.10(.85)	2.17(.86)	181
Prior belief not supporting the topic	Prior belief compatible with text	2.49(.96)	2.19(.90)	75
	Prior belief incompatible with text	1.86(.76)	2.05(.83)	75

Descriptive were examined on the relationship between three interactive sources of

cognitive clashes: prior belief incongruence with the position of the text, agent disagreement,

agent's position, and incongruence with the position of the text, participants had the highest

summary scores when they experienced no clash ($M = 2.50$, $SD = .84$). Participants' summary

score was lowest when their prior beliefs were congruent with the text and the agents disagreed

with each other, with the teacher having the position that's compatible with the text ($M = 2.00$,

$SD = .77$).

As a follow-up analysis, separate models were constructed either using only the data with

strong prior belief ratings (i.e. ratings of 1 and 2 for not support, ratings of 5 and 6 for support),

or the data with neutral ratings (i.e. ratings of 3 for not support, ratings of 4 for support). There

was a significant main effect of participants' prior belief compatibility with the position of the

text ($F(1, 218.3) = 16.68$, $p < .001$, $MSe = 8.51$, $\eta^2 = .05$). Contrary to the results found for

participants' sentence recognition for meaning, when participants reported strong and solid

opinions of the topic, the participants who reported their prior beliefs compatible with the

position of the text had significantly higher summary scores ($M = 2.28$, $SD = .91$, $N = 140$) than

the participants who reported their prior beliefs were incompatible with the position of the text

($M = 1.90$, $SD = .76$, $N = 136$). There was also a significant main effect of participants' prior

belief compatibility with the position of the text on participants' summary scores ($F(1, 210.48) =$

3.88, $p = .05$, $MSe = 1.38$, $\eta^2 = .01$) when only using the data with more neutral prior belief

ratings (i.e. ratings of 3 for not support, ratings of 4 for support). Albeit less strong of an effect,

when participants reported more neutral stance on the topic, the participants who reported their

prior beliefs compatible with the position of the text also had significantly higher summary

scores ($M = 2.28$, $SD = .88$, $N = 120$) than the participants who reported their prior beliefs

incompatible with the position of the text, ($M = 2.09$, $SD = .79$, $N = 116$). Replicating the previous findings, there was a significant interaction between participants compatibility of prior belief to the position of text and whether the agents agreed with each other ($F(1, 195.71) = 4.58$, $p = 0.03$, $MSe = 2.02$, $\eta^2 = .001$). When participants listened to the agents disagreeing with each other, the participants who reported their prior beliefs incompatible with the position of the text had significantly higher summary scores ($M = 2.39$, $SD = .91$, $N = 60$) than participants who reported their prior beliefs were compatible with the position of the text ($M = 1.98$, $SD = .67$, $N = 61$; $t(193.70) = 2.71$, $p = .04$).

When topic was added into the model as a fixed factor, there was a significant main effect of topic on the participants' summary scores of the text ($F(3, 381) = 4.48$, $p = .004$, $MSe = 2.18$, $\eta^2 = 0.02$). A post hoc test revealed that, consistent with the result on participants' sentence recognition for meaning, the topic regarding the Black Lives Matter Movement had significantly lower average summary scores ($M = 1.98$, $SD = .80$, $N = 128$) than either the texts with the topic of the Men's Rights Movement ($M = 2.26$, $SD = .88$, $N = 128$; $t(381) = -3.14$, $p = .01$) or the Syrian Refugee Crisis ($M = 2.23$, $SD = .83$, $N = 128$; $t(381) = -2.82$, $p = .03$), whereas the summary scores between the topic on Black Lives Matter Movement and the Men's Rights Movement, Black Lives Matter Movement and the Trump Presidency did not significantly differ from each other. The topics on Men's Rights Movement, the Trump Presidency, and the Syrian Refugee Crisis also did not differ significantly from each other in terms of participants' summary scores.

Four separate models were conducted for each of the topics. Only the topic on the Men's Rights Movement and the Trump Presidency yielded significant models on any of the independent variables. Contrary to the analysis conducted on participants' recognition for

meaning, the results show that only when the participants read texts regarding the Men's Rights Movement and the Trump Presidency was there a significant main effect of compatibility of text position with prior beliefs on the participants' summary scores ($F(1, 126) = 8.14$, $p = .005$, $MSe = 5.98$, $\eta^2 = 0.06$; $F(1, 126) = 9.18$, $p = .003$, $MSe = 6.48$, $\eta^2 = 0.07$, respectively).

For the topic on Men's Rights Movement, consistent with the previous analysis on recognition for meaning, when participants' prior beliefs were reported to be incompatible with the position of the texts on the topic of the Men's Rights Movement, they had a significantly higher summary score ($M = 2.49$, $SD = .86$, $N = 59$) than when their prior belief was compatible with the position of the ($M = 2.06$, $SD = .86$, $N = 69$). There was a significant interaction between whether the agents agreed with each other and participants' prior beliefs on the position of the topic affecting participants' recognition memory for the original meaning on the topic of the Men's Rights Movement ($F(1, 124) = 9.07$, $p < .001$). While reading texts on the topic of the Men's Rights Movement, when participants encountered agents disagreeing with each other, the participants who reported their prior beliefs incompatible with the position of the text had significantly higher summary scores ($M = 2.72$, $SD = .84$, $N = 29$) than participants who reported their prior beliefs compatible with the position of the text ($M = 1.76$, $SD = .76$, $N = 29$; $t(124) = 4.707$, $p < .001$).

On the other hand, on the topic of the Trump Presidency, contrary to previous findings, when the participants reported their prior beliefs were compatible with the position of the text, they had significantly higher summary scores ($M = 2.31$, $SD = .91$, $N = 59$) than the participants who reported their prior beliefs incompatible with the position of the text ($M = 1.86$, $SD = .78$, $N = 69$). There was no interaction between participants' prior beliefs and the agent speeches for the topic of the Trump Presidency.

51

Source memory scores. On average, participants correctly identified 0.49 (*SD* = .50) of agent sources on the question items, which was significantly above chance level of 0.33 (*t*(8191) = 29.23, *p* < .001). We first investigated whether participants could identify old items from new items. A mixed-effects logistic regression model for the correct (coded as 1) or incorrect (coded as 0) source identification between old and new items yielded a significantly better fit than did a model with only the random effect but without the fixed effects, (χ^2(1) = 167.99, *p* < .001). The identification of new items was 1.81 ($e^{0.59}$) times better than the identification of old items among the participants.

Summary descriptive statistics for participants' source memory are presented in Table 5. Based on the proportion scores, we found that there was a slight response bias towards guessing the statement as a new item when the statement was in fact spoken by the teacher agent. There was also a slight bias towards guessing the student agent as the source when the statement was in fact a new item. On the other hand, when the statement was spoken by the student agent, there was no difference between guessing the teacher agent and guessing that it was a new item. Furthermore, correct rejection (i.e. correctly responding the item was new when in fact it was) had the greatest proportion of hits.

Table 5. Proportion of agent source responses.

	Participant Response		
Agent Source	Teacher	Student	Neither
Teacher	0.42	0.27	0.31
Student	0.29	0.42	0.29
Neither	0.21	0.23	0.56

A mixed-effects logistic regression model for source identification on presented speech statements did not yield a significantly better fit than did a model with only the random effect but without the fixed effects, ($\chi^2(1) = 2.38$, $p = .12$). However, the Wald Z-test using assumption of normal distribution likelihood yielded significance for whether the agents agreed or disagreed with each other as a fixed factor ($Z = 2.25$, $p = .02$). Summary descriptive statistics for participants' source memory per condition are presented in Table 6.

When participants listened to the agents agreeing with each other, they had significantly better source memory ($M = 0.50$, $SD = 0.50$) than when they listened to the agents disagreeing with each other ($M = 0.48$, $SD = 0.50$). There was also a significant main effect of the agent speech conditions T-S- and T+S- ($Z = -2.47$, $p = .01$). When agents agreed in opposition to the text position, participants had a significantly better source memory ($M = 0.51$, $SD = 0.50$) than when the agents disagreed with each other but the teacher's position was compatible with the position of the text ($M = 0.47$, $SD = 0.50$). For participants' confidence of identifying the correct source, a model with both fixed and random effect resulted in a better fit than a random effect only model ($\chi^2(1) = 149.89$, $p < .001$). The participants had a source memory increase of 1.25 ($e^{0.22}$) times when they rated higher confidence on identifying the source. There was no significant main effect between the other agent speech conditions. It was also found that there was no topic or prior belief effects.

Table 6. Descriptive statistics for discrimination memory for source per condition.

Prior Belief Conditions		Agent Speech Conditions		
		Agents agreeing	Agents disagreeing	N Total
Prior belief supporting the topic	Prior belief compatible with text	0.55(.50)	0.51(.50)	181
	Prior belief incompatible with text	0.55(.50)	0.37(.49)	181
Prior belief not supporting the topic	Prior belief compatible with text	0.41(.50)	0.53(.51)	75
	Prior belief incompatible with text	0.60(.50)	0.45(.51)	75

The source memory exam was given to the participants at the very end of the experiment. It is possible that there was a decrease in source memory due to cumulated fatigue. To investigate this further, a mixed-effects logistic regression model was conducted on the correct and incorrect of source identification on just the first pages of each text. The resulting model yielded a significantly better fit than a model with only the random effect but without the fixed effects, ($\chi^2(1) = 4.10$, $p = .04$). Furthermore, the model replicated the findings of the model with all data. When participants listened to the agents agreeing with each other, they had significantly better source memory ($M = 0.50$, $SD = 0.50$) than when they listened to the agents disagreeing with each other ($M = 0.46$, $SD = 0.50$). In other words, the participants had a source memory increase of 1.20 ($e^{0.18}$) times when they listened to the agents agreeing with each other than when they listened to the agents disagreeing with each other. There was no significant difference of source memory between when the teacher's position is compatible versus incompatible with the position of text.

Separate mixed-effects logistic regression models were conducted for each of the agent sources (teacher, student, neither) with respect to the independent variables—topic support, prior belief compatibility to text position, and agent agreement/disagreement—on the correct and incorrect of source identification. Full summary descriptive statistics for participants' source memory per condition per agent source are presented in Table 7-9. For the statements spoken by the teacher agent, the resulting model yielded a significantly better fit than a model with only the random effect but without the fixed effects, when only prior belief compatibility to text position was the fixed effect ($\chi^2(1) = 4.89$, $p = .03$). The participants who reported their prior beliefs compatible with the position of the text had significantly better source memory ($M = 0.45$, $SD = 0.50$) than participants who reported their prior beliefs incompatible with the position of the text ($M = 0.39$, $SD = 0.49$). In other words, the participants had a source memory increase of 1.45 ($e^{0.37}$) times when their prior beliefs were compatible with the position of the text when the source of the speech statement was the teacher agent. For the statements spoken by the student agent, the resulting model did not yield a significantly better fit than did a model with only the random effect but without the fixed effects on any of the independent variables ($\chi^2(1) = 0.37$, $p = .95$). Finally, for the statements that were new (not spoken by either agents), the resulting model yielded a significantly better fit than a model with only the random effect but without the fixed effects, for both topic support and agent agreement/disagreements, but not prior belief compatibility to text position ($\chi^2(1) = 5.52$, $p = .02$ and $\chi^2(1) = 16.82$, $p < .001$, respectively). There were no interaction effects. The participants who reported not supporting the topic had significantly better source memory (($M = 0.59$, $SD = 0.49$) than participants who reported supporting the topic ($M = 0.55$, $SD = 0.50$). In other words, the participants had a source memory increase of 1.21 ($e^{0.19}$) times when they reported not supporting the topic when the statement

55

presented to them were new items. When participants listened to the agents agreeing with each other, they had significantly better source memory ($M = 0.59$, $SD = 0.49$) than when they listened to the agents disagreeing with each other ($M = 0.53$, $SD = 0.50$). In other words, the participants had a source memory increase of 1.46 ($e^{0.38}$) times when they listened to the agents agreeing with each other than when they listened to the agents disagreeing with each other when faced with new items.

Table 7. Descriptive statistics for discrimination memory for source per condition on the statements spoken by the teacher agent.

Prior Belief	Agent Speech	Response	Proportion
Supporting the topic	Compatible with text — Agreeing	Teacher	0.43
		Student	0.29
		Neither	0.29
	Compatible with text — Disagreeing	Teacher	0.49
		Student	0.25
		Neither	0.26
	Incompatible with text — Agreeing	Teacher	0.41
		Student	0.29
		Neither	0.30
	Incompatible with text — Disagreeing	Teacher	0.37
		Student	0.26
		Neither	0.37
Not supporting the topic	Compatible with text — Agreeing	Teacher	0.39
		Student	0.31
		Neither	0.31
	Compatible with text — Disagreeing	Teacher	0.45
		Student	0.26
		Neither	0.28
	Incompatible with text — Agreeing	Teacher	0.37
		Student	0.32
		Neither	0.31
	Incompatible with text — Disagreeing	Teacher	0.44
		Student	0.23
		Neither	0.33

Table 8. Descriptive statistics for discrimination memory for source per condition on the statements spoken by the student agent.

Prior Belief		Agent Speech	Response	Proportion
Supporting the topic	Compatible with text	Agreeing	Student	0.45
			Teacher	0.27
			Neither	0.28
		Disagreeing	Student	0.44
			Teacher	0.28
			Neither	0.28
	Incompatible with text	Agreeing	Student	0.42
			Teacher	0.33
			Neither	0.25
		Disagreeing	Student	0.42
			Teacher	0.26
			Neither	0.33
Not supporting the topic	Compatible with text	Agreeing	Student	0.38
			Teacher	0.32
			Neither	0.30
		Disagreeing	Student	0.37
			Teacher	0.27
			Neither	0.36
	Incompatible with text	Agreeing	Student	0.39
			Teacher	0.35
			Neither	0.26
		Disagreeing	Student	0.48
			Teacher	0.24
			Neither	0.29

Table 9. Descriptive statistics for discrimination memory for source per condition on the new statements.

Prior Belief		Agent Speech	Response	Proportion
Supporting the topic	Compatible with text	Agreeing	Neither	0.59
			Teacher	0.20
			Student	0.21
		Disagreeing	Neither	0.50
			Teacher	0.25
			Student	0.25
	Incompatible with text	Agreeing	Neither	0.59
			Teacher	0.16
			Student	0.25
		Disagreeing	Neither	0.52
			Teacher	0.25
			Student	0.24
Not supporting the topic	Compatible with text	Agreeing	Neither	0.57
			Teacher	0.18
			Student	0.25
		Disagreeing	Neither	0.62
			Teacher	0.21
			Student	0.18
	Incompatible with text	Agreeing	Neither	0.59
			Teacher	0.21
			Student	0.21
		Disagreeing	Neither	0.31
			Teacher	0.37
			Student	0.32

Source memory was included as a mediator, which was estimated in multi-level mediation models (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014). Average causal estimates were estimated using a quasi-Bayesian Monte Carlo method with 10,000 simulations. This estimation approach, detailed in Imai, Keele, and Tingley (2010), provided point estimates, 95% confidence intervals, and p-values for the mediated, direct, and total effects. The mediated effect

would be the average effect of the fixed factors on the summary scores and discrimination scores that was due to the mediator alone. The direct effect can be interpreted as the average effect of the fixed factors not due to the mediator and the total effect is the average combination of the mediated and direct effects. Two model sets were specified to assess the average causal effect for each dependent variable: a mediator model (predicting the mediator) and an outcome variable model (predicting summary scores and discrimination scores for both surface structure and meaning). Therefore, a total of 3 model sets were conducted for the three dependent variables. Since affect (valence/arousal) was never found to be a significant covariate, the two variables were dropped from the mediation model in the models using the quasi-Bayesian Monte Carlo simulation method. Using the "mediation" package in R, the models with source memory as a mediating factor to predict summary scores and both discrimination scores for surface structure and meaning was never significant. Indeed, the resulting models show that source memory had 0 mediating effect on either surface structure or meaning.

Memory detection, agent source discrimination, and guessing for teacher and student estimates were computed using proportion scores. An "old item" on the source memory exam is a statement that was spoken by one of the agents. A "new item" on the source memory exam is a statement that was not spoken by any agent. The proportion of hits, misses, false alarm, and correct rejection were computed as the following:

Hits = The proportion of selecting the correct agent source on an old item.
Misses = The proportion of choosing "Neither" on an old item.
False Alarm = The proportion of choosing an agent source on a new item.
Correct Rejection = The proportion of choose "Neither" on a new item.

Based on the calculation, the proportion score for the hits was 0.42, the proportion score for the misses was 0.30, the proportion score for the false alarm was 0.44, and the proportion score for the correct rejection was 0.56.

The probability of memory detection (detection of old item) was computed as the proportion of choosing any agent on an old item. The probability of agent source discrimination was computed as the proportion of selecting the correct agent source out of the total proportion of correct responses. The probability of guessing teacher was computed as the proportion of choosing the teacher agent out of the total proportion of false alarms. The probability of guessing student was computed as the proportion of choosing the student agent out of the total proportion of false alarms. From these computations, memory detection was a score of 0.70, agent discrimination was a score of 0.43, guessing teacher was a score of 0.48, and guessing student was a score of 0.52.

In order to account for response bias and hits from guessing, three General Tree Processing (GTP) models (Riefer & Batchelder, 1995) were constructed to assess the participants' detection, discrimination, and guessing for teacher, student, and neither (new items). Figure 5 shows all three source models constructed for teacher, student, and new. The models assessed whether the participants have a bias towards the teach or student agent through guessing. The models assumed that the probabilities of the memory response categories (i.e., teacher agent, student agent, neither) are a function of four hypothetical (latent) parameters: D for detection of old items, t for discrimination between sources, B for bias, and G for guessing. There are only three sources, so a saturated model was used with 5 parameters with 6 degrees of freedom, where only one type of guessing was used. In this decision tree model, each source model started with parameter D for detection. The probability of the parameter D was then tested

against the probability of 1- D for the failure of detection. If detection (D) is successful, then

discrimination t was tested against 1-t for failure of discrimination. If discrimination was

successful, it was assumed that the participant recalls the old item as well as being able to

successfully discriminate who said what. If, however, the participant failed to discriminate, then

a guessing parameter G for the sources was estimated. If on the other hand the participant failed

to detect the old item to begin with, then a bias parameter B was to be tested against 1-B for lack

of bias. If the result shows that bias does indeed exist, the guessing G parameter was estimated.

If, however, a bias was not present, then the model would indicate that the participant had judged

the item to be a new item. The full equation of the saturated model is shown below:

Teacher detection: $Dt*td + Dt*(1-td)*G + (1-Dt)*B*G$

Teacher discrimination: $(1-Dt)*B*(1-G) + Dt*(1-td)*(1-G)$

Teacher guess: $(1-Dt)*(1-B)$

Student detection: $Ds*ts + Ds*(1-ts)*(1-G) + (1-Ds)*B*(1-G)$

Student discrimination: $(1-Ds)*B*G + Ds*(1-ts)*G$

Student guess: $(1-Ds)*(1-B)$

Biased guess new item: $B*G$

Bias no guess new item: $B*(1-G)$

Detection of new item: $(1-B)$

The resulting model using the 5-parameter saturated model was a near-perfect fit to the

data, with a difference of log-likelihood of 1^{E-10}. The detection probability between the source of

teacher and the source of student was tested. The result showed that there was no difference in

detection, discrimination or guessing between the source of teacher (estimate of 0.45 for

detection, 0.37 for discrimination, and 0.48 for guessing) and student (estimate of 0.49 for

detection, 0.23 for discrimination, and 0.52 for guessing). However there seemed to be a slight

bias for student (estimate of 0.57). When a Monte Carlo simulation was conducted, the resulting

parameters indicated that to have best detection for the current model, an estimate of at least 0.80

must be reached for the D parameters for the model to have at least 95% probability to reject the

null hypothesis. This suggests that the model with the current data was under powered. A larger

sample size or items with better discriminability may achieve better results.
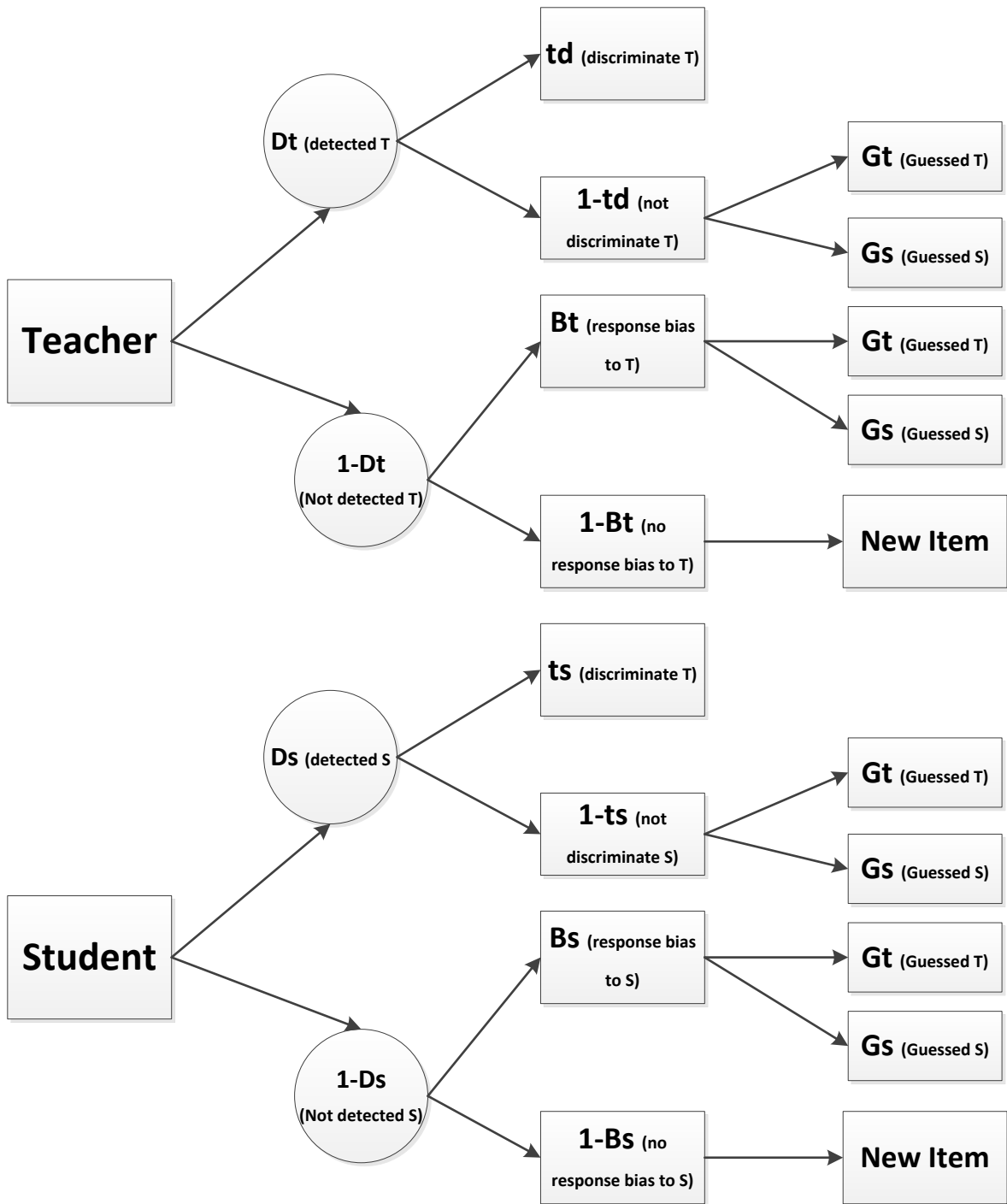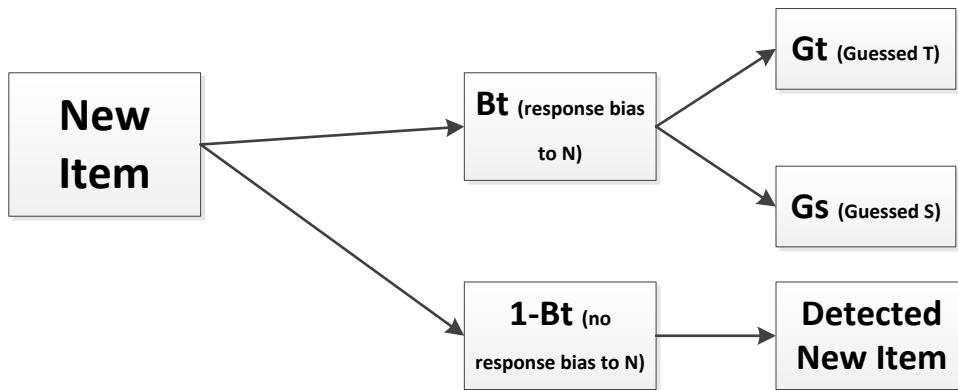
Figure 5. General Processing Tree Models (the Bs and Gs are one parameter).

Reading times. Reading time per page was recorded in seconds. Due to the variation of the text lengths per page, an average reading time per 100 words was computed for each page. On average, the participants spent 31.68 seconds per 100 words per page (*SD* = 16.69). Summary descriptive statistics for participants' reading time per 100 words per page are presented in Table 10.

There was a significant main effect of compatibility of text position with prior beliefs on the participants' reading time (*F*(1, 1978.3) = 18.30. *p* < .001, *MSe* = 2849.41, η^2 = .006). When the participants reported their prior beliefs congruent with the position of the text, the participants spent significant longer time reading the pages (*M* = 32.91, *SD* = 18.07, *N* = 1024) than when they reported their prior beliefs incongruent with the position of the text (*M* = 30.44, *SD* = 15.09, *N* = 1024). There was a significant interaction effect between the participants' prior beliefs on the position of topic and participants' compatibility of text position with prior beliefs on participants' reading time (*F*(1, 1931.7) = 16.71, *p* < .001, *MSe* = 2560.78, η^2 = .005). When participants' prior beliefs on the position of the topic supported the topic, participants who had compatible prior beliefs with the position of the text had significantly longer reading time (*M* =

34.53, $SD = 18.91$, $N = 724$) than participants who had incompatible prior beliefs with the position of the text ($M = 30.57$, $SD = 15.24$, $N = 724$; $t(1942.32) = 5.58$, $p < .001$). Furthermore, when both participants' prior beliefs on the position of the topic support the topic, and their prior belief was compatible with the position of the text, there was also an increased reading time ($M = 34.53$, $SD = 18.91$, $N = 724$) relative to participants who both had prior belief not supporting the topic and incompatible with the position of the text ($M = 30.12$, $SD = 14.74$, $N = 300$; $t(1948.36) = 3.97$, $p < .001$). There was a second interaction effect between compatibility of text position with prior beliefs and whether agents agreed with each other on participants' reading time ($F(1, 463.89) = 7.44$. $p = .007$, $MSe = 3.65$, $\eta^2 = .02$). When the participants reported their prior beliefs congruent with the position of the text, the participants who listened to the agents disagreeing with each other had a significantly longer reading time ($M = 34.80$, $SD = 19.00$, $N = 524$) than the participants who listened to the agents agreeing with each other ($M = 30.94.02$, $SD = 16.83$, $N = 125$; $t(1985.73) = 3.02$, $p = .01$).

Table 10. Descriptive statistics for reading time per 100 words per page.

Prior Belief Conditions		Agent Speech Conditions		
		Agents agreeing	Agents disagreeing	N Total
Prior belief supporting the topic	Prior belief compatible with text	32.44(16.57)	36.43(20.63)	724
	Prior belief incompatible with text	31.95(16.35)	29.23(13.98)	724
Prior belief not supporting the topic	Prior belief compatible with text	27.63(16.98)	30.49(12.90)	300
	Prior belief incompatible with text	30.62(16.23)	29.48(12.63)	300

Descriptive statistics were examined on the relationship between the three interactive sources of cognitive clashes: 1. prior belief incongruent with the position of the text, 2. agent disagreement, and 3. agent's position incongruent with the position of the text. Participants had the longest reading time when they only experienced clash of agent disagreement, where the teacher had the position that's compatible with the position of the text ($M = 35.09$, $SD = 19.99$). On the other hand, the participants' reading time was the shortest when their prior belief was incompatible with the position of the text and the agents also disagree with each other (teacher compatible: $M = 29.25$, $SD = 14.14$; teacher incompatible: $M = 29.34$, $SD = 13.14$).

In addition, when separate models were constructed either using only the data with extreme prior belief ratings (i.e. ratings of 1 and 2 for not support, ratings of 5 and 6 for support) or more neutral prior belief ratings (i.e. ratings of 3 for not support, ratings of 4 for support), the models on reading time replicated the models on participants' recognition memory for surface structure. There was a significant interaction effect of participants' prior belief support for topic and prior belief compatibility with the position of the text, but there were no main effects or interaction effects of the agent speech conditions ($F(1, 1083.5) = 8.24$, $p = .004$, $MSe = 1194.50$, $\eta^2 = .002$). When participants' prior beliefs strongly supported the topic, the participants who reported their prior beliefs compatible with the position of the text had significantly longer reading times ($M = 32.93$, $SD = 17.20$, $N = 352$) than participants who reported their prior beliefs incongruent with the position of the text ($M = 30.09$, $SD = 15.23$, $N = 352$; $t(1056.53) = 3.53$, $p = .003$).

On the other hand, when only using the data with more neutral prior belief ratings (i.e. ratings of 3 for not support, ratings of 4 for support), there was a significant interaction effect between participants' compatibility of prior belief with the position of the text and whether the

agents agreed or disagreed with each other ($F(1, 927.65) = 11.12$, $p < .001$, $MSe = 1523.70$, $\eta^2 = .02$). When participants prior beliefs are compatible with the position of the text but had a more neutral position on the topic, they had a significantly longer reading time when they listened to the agents disagreeing with each other ($M = 36.59$, $SD = 21.03$, $N = 244$) than the participants who listened to the agents agreeing with each other ($M = 31.65$, $SD = 16.96$, $N = 220$; $t(902.60) = 3.10$, $p = .01$).

When topic was added into the model as a fixed factor, there was a significant main effect of topic on the participants' reading time per page ($F(3, 1920.3) = 3.52$, $p = .01$, $MSe = 537.11$, $\eta^2 = .003$) and a post hoc test revealed that only the topic regarding the Trump Presidency and the Syrian Refugee Crisis significantly differ in reading time. Participants spent much shorter time reading the texts on the Trump Presidency ($M = 30.20$, $SD = 15.47$, $N = 512$) than the texts on the Syrian Refugee Crisis ($M = 32.77$, $SD = 17.27$, $N = 512$; $t(1933.61) = -3.10$, $p = .01$). Four separate models were conducted for each of the topics. The results show that only the topic on the Trump Presidency did not yield significant models on any of the independent variables. The topics Black Lives Matter and The Syrian Refugee Crisis yielded significant interactions between participants' compatibility of prior beliefs to the position of the text ($F(1, 298.94) = 8.00$, $p = .005$, $MSe = 2391.32$, $\eta^2 = 0.02$; $F(1, 275.4) = 21.88$, $p < .001$, $MSe = 6027.5$, $\eta^2 = 0.04$, respectively). However, post hoc tests indicated that only the topic on the Syrian Refugee Crisis showed significant differences in reading times. When their prior belief was compatible with the position of the text but listened to the agents disagreeing with each other, participants spent significantly longer time reading the texts ($M = 40.83$, $SD = 18.54$, $N = 124$) than when their prior belief was compatible with the position of the text but listened to the agents agreed with each other ($M = 30.65$, $SD = 17.33$, $N = 132$; $t(504) = 3.99$, $p < .001$). For the topic

on the Men's Rights Movement, there was a significant main effect on prior belief support for topic ($F(1, 504) = 11.77$, $p < .001$, $MSe = 3030.51$, $\eta^2 = 0.0009$). Consistent with the previous analysis, when participants' prior beliefs supported the topic, they spent significantly longer time reading the texts on the topic of the Men's Rights Movement ($M = 32.58$, $SD = 16.90$, $N = 428$) than when participants' prior beliefs did not support the topic ($M = 26.01$, $SD = 12.10$, $N = 84$). Finally, for the topic of the Men's Rights Movement, there was a significant main effect for the compatibility of participants' prior beliefs to the position of the text ($F(1, 504) = 7.35$, $p = .006$, $MSe = 1892.18$, $\eta^2 = 0.00005$). When participants' prior beliefs were reported to be compatible with the position of the texts on the topic of the Men's Rights Movement, they spent significantly longer time reading the texts ($M = 33.12$, $SD = 17.86$, $N = 276$) than when their prior belief was incompatible with the position of the texts ($M = 29.62$, $SD = 14.28$, $N = 236$). On the topic of the Men's Rights Movement there was no significant main effect or interaction effect with the agent speech conditions.

Discussion

Overview of Research

The current theory in cognitive and learning sciences on memory and comprehension has established the importance of both the role of cognitive disequilibrium (D'Mello et al., 2015; Graesser et al., 2015) and source monitoring as a memory retrieval strategy to potentially resolve conflicting information during reading (Braasch et al., 2012; Goldman et al., 2012; Rouet, 2006). However, a serious study has yet to be conducted that combines these two major lines of research to investigate learning texts on controversial topics while listening to agents disagreeing with each other. Texts with controversial topics are of great interest and amply relevant to the current times because they are "hot topics" with conflicting viewpoints which could drive

learners to be emotionally invested into comprehending the texts. In addition, participants' prior beliefs on controversial topics would create the condition to induce them into a state of cognitive disequilibrium if the prior belief is incongruent with the position of the text.

Summary of the Results

This dissertation investigated the memory and learning of texts and spoken messages after participants read controversial texts with different positions and listened to animated agents discussing the texts with different positions. The differential positions were manipulated to examine the relationships between the different sources of clashes: participants' prior belief clashing with the position of the text, prior belief clashing with agent statements, agent position clashing with text position, and agent disagreements with each other. In addition, this dissertation investigated source monitoring as one potential strategy to resolve the conflict stimulated from these clashes of cognitive disequilibrium. We used *AutoTutor* (Graesser, 2016; Graesser et al., 2016), an artificially simulated learning environment with two agents to model such a complexly dynamic human learning from text. We hypothesized that participants would have better discriminatory recognition memory for texts with positions that are incompatible with their prior beliefs. On the other hand, we hypothesized that participants would have better recall memory on texts with positions that are compatible with the participants' prior beliefs. Our results supported both hypotheses. Pertaining to memory, we found that overall, participants' prior beliefs was the largest predictor of both recognition and recall memory through summarization. The surprising result was that the agents having a discussion back and forth regarding the text barely factored into memory for these controversial texts. In fact, our results show that agents only mattered for texts that the participants have a neutral stance on the topic (somewhat support or somewhat not support). Listening to agent disagreements predicted better memory for the participants, but only

for memory recall and recognition memory for surface structure, not recognition for meaning. This suggests that when participants have not made up their mind about an issue or perhaps know very little about the issue, listening to a discussion with contrary view points helped their comprehension and memory. When participants have made up their minds about an issue, agents having contrary opinions does not help with either memory or comprehension of the text.

The current results on source memory for the agent speeches replicated previous findings that participants have a stronger memory detection for the old items when the agents agreed with each other (Feng et al., 2017). Another replication from previous study is that, as revealed from the general tree models, the participants could not robustly discriminate the agent sources. However, a rudimentary signal-detection computation did suggest that some discrimination had occurred at better than chance levels. In addition, no mediating effect of source memory to text memory or summarization was found. This is possibly due to the source memory test being administered at the very end of the experiment so there was an effect of *retrieval induced forgetting* (Anderson, Bjork, & Bjork, 2000; Storm, Bjork, Bjork, & Nestojko, 2006), where attending to texts for later retrieval resulted in forgetting the agent speeches associated with the texts. It is possible that the participants either did not put much effort in answering the questions or that memory by that time has waned to the level of guessing. Indeed, when the analysis was conducted on only the first page of each text, the resulting model was a significantly better fit than the model with only random effects. The analysis overall still supports the notion that the most salient source memory for agent speeches occurred when the participants listened to the agents agreeing in opposition to the position of the text. This once again suggests that cognitive disequilibrium is still at work in providing saliency of information.

Not only so, the effect of agent positions still seemed to only matter when comparison was between the position of the teacher agent and the position of the text. In other words, when the teacher agent's position contradicted the position of the text, and this position is fortified by the agreement of the student agent, the participants can better remember and identify the source of information. This is further supported by the finding that participants had higher accuracy in sentence recognition only when the teacher agent's position was incompatible with the position of the text. This is especially the case for both the recognition of surface structure and summary scores. These results also replicate the results from previous research that showed a bias in the learner attending to the teacher (Feng et al., 2017). There was greater impact of the teacher as a potential reliable source of disseminated information than the student or peer.

Finally, the current study found that the topic of the texts was a significant factor in participants' memory and comprehension. The texts on the topic of the Syrian Refugee Crisis seemed to have particularly mediated the effect of the agent speech conditions. The participants rated the texts on the topic of the Syrian Refugee Crisis to be the least emotionally arousing. These results seem to support the finding that the agent disagreements have a significant effect on the participants' memory for text when the participants reported their prior belief to be more neutral in terms of supporting the position of the topic, even though the ratings of affect are not significantly different between the topics. It has also been found that when the participants read texts on the Trump Presidency, they had better recognition memory for surface structure than all the other topics. This is supported by Long (1989), who reported that reading emotionally charged texts would result in better surface structure of the text.

Participants reported their prior belief to be significantly less supportive of the topic of the Trump Presidency than on the topics of Black Lives Matter, Men's Rights Movement, and

the Syrian Refugee Crisis, whereas the latter three topics did not significantly differ in ratings among each other. Past research has demonstrated that negative emotional content resulted in better long-term memory retention than positive or neutral content, especially for delayed recall tasks (e.g. Kensinger, 2007). In addition, despite the non-significant differences, the topic of the Trump Presidency was rated the highest on emotional valence, but not the highest on arousal. The affect ratings themselves did not seem to influence as a co-variate on any of the fixed effects, despite the finding that higher valence and higher arousal were both found for prior belief compatible texts. These results suggest that the topic mediated the effects of participants' prior beliefs and the agent speech conditions independent from participants' own ratings of affect and even prior beliefs.

The current study has supported findings from previous research as well as combined the theories together in the dynamic system of memory for textual information discussed by agents. According to the *Schema Pointer + Tag* hypothesis, readers should have better recognition or discriminatory memory for atypical information (Graesser & Nakamura, 1982). The current study showed that prior belief is the single best predictor of recognition memory for text. When participants read texts with positions that are incompatible with their prior belief, they had higher recognition memory for both the surface structure and the meaning. This suggests that the texts that contain information that violated the participants' schema regarding the topic were atypical to the participants, which resulted in better discriminatory memory for the participants. Several memory models also support the finding that atypical information is more likely to be discriminated from typical information during memory recognition rather than memory recall. Hunt and McDaniel (1993) suggest that while semantically similar items may facilitate information generation in memory recall, the greater learning of semantically similar items may

come at the expense of memory discrimination, which is better for semantically unrelated items. Similarly, Raaijmakers and Shiffrin's (1981) *Semantic Association Model* also suggests that items having high similarity would result in a stronger network of semantic associations. Therefore, prior belief congruent information can lead to superior generation of those items in memory recall but lower memory discrimination between individual items.

Stangor and McMillan (1992) conducted a meta-analysis of 54 experiments that investigated the influence of social expectations on memory for information that is either congruent or incongruent with those expectations. Collectively, they found that stronger expectations would lead to increased schematic processing, and therefore better memory, for expectancy-congruent information, whereas recognition sensitive information that is expectation incongruent would be processed in a more salient, individualistic manner than expectancy-incongruent information. However, Stangor and McMillan's (1992) meta-analysis showed that both memory recall and memory recognition tend to be overall better for expectancy-incongruent information. They found that recognition memory benefited more than recall on expectancy-incongruent information when the expectancy was better developed in the participants' schemas. One explanation for the current results is that because participants' summaries reflected their comprehension of the arguments presented in the text rather than pure memory recall of the information. As demonstrated by previous research on argumentation, in order to generate accurate information and articulate arguments from the texts, coherent situation model would need to be constructed (Voss & Wiley, 1999). Prior belief congruency would assist in constructing that coherent situation model. The schematic processing coincided with the expectancy-congruent information presented in the text would therefore produce better summary than expectancy-incongruent information. Not only so, the nature of the controversial topic may

produce higher expectancy than other more generic topics in which memory recall did not benefit from incongruence.

The results in reading time further provided evidence regarding the relationship between prior beliefs and encountering agent disagreement on participants' memory measures. The participants took a longer amount of time while reading the text with position that was congruent with their prior beliefs, suggesting better engagement for those texts. On the other hand, when participants read texts in which they did not have a strong stance on the topic, they took a longer amount of time reading the text when they encountered agent disagreements. This suggests that agent disagreement better stimulated cognitive disequilibrium when the participants did not have a strong schema built around the issues that were presented, and could be swayed either way; therefore, a longer amount of time was needed to construct the mental model and to resolve the conflict.

The current study formulated three competing hypotheses on stimulating cognitive disequilibrium to attend to agent sources: 1. the either/or hypothesis—participants attend to sources either when the agents contradict each other through disagreements or when there was an incongruence between position of text and participants prior beliefs, 2. the additive hypothesis—both clashes would increment the effect, and 3. the both-and hypothesis–only when both clashes occur together will participants attend to sources. Surprisingly, our results did not support any of the three hypotheses, as source memory did not seem to mediate memory or comprehension for text. Interestingly, even information espoused by the agents that would contradict participants' prior belief did not seem to have an effect on their source memory. This result seemed to contradict previous findings which suggested that participants are more likely to resort to source monitoring when they read texts that contradict their prior beliefs (Braten et al., 2016).

One explanation could be that since the participants were constantly exposed to contrary sides from the agents who, in turn, can contradict the text, the participants represented the agent information separately from the text and used the text as the reference to tag atypical information espoused from the agents. Participants can discriminate the sources better only when both agents agree in contradicting the text. One explanation might be that the participants' cognitive resources had been largely used up in an effort to comprehend the text and build a successful situation model; therefore, there was little resource left to also remember the speeches. On the other hand, when participants did not have a solid schema for the topic, as reflected by their neutral ratings of topic support, the agents seemed to help the participants in recognizing sentence structures and meaning as well as in retrieving information for memory recall and summarization.

The most logical explanation is that since the source memory test was given at the end of the experiment, the participants' memory for agent speeches had waned over time. In addition, there were not enough cognitive resources for the participants to remember both the agent speeches and the texts. The fact that the speech statements were presented to the participant only one time, whereas the participants were able to re-read any parts of the text within a single page, would reasonably make it more difficult for the participants to remember the speeches relative to the text. This is also shown by the general tree model (Riefer & Batchelder, 1995) which suggested that overall participants could not robustly recall or discriminate between the sources due to high guessing parameters. On the other hand, consistent with the previous research that suggests that conflict and contradictions will trigger more source monitoring from participants (Braasch et al., 2012; Rouet, 2006), the current study still found that a clash did predict higher source memory. Our results showed that the only clash that seemed to result in better source

memory was when the agents agreed with each other, but both contradicted the position of the text. This agent speech condition resulted in better identification of agent source when compared to the condition where the agents disagreed and the teacher agent's statement was compatible with the position of the text, which once again suggests the high weight of the teacher as a source of information (Rourke & Anderson, 2002).

The current result replicated Feng et al.'s (2017) findings, which also suggested that agent source memory did not predict comprehension even when the agents were speaking a sentence that was verbatim from the text. Other than memory waning resulting in the poor performance in identifying source, it is also possible that contrary to the previous source-monitoring research, the agent sources are not embedded in the text and therefore may be stored in separate memory representations from the text rather than linked together as described in the *document model framework* (Britt & Rouet, 2012) and the *discrepancy-induced source comprehension* model (Braasch et al. 2012). The lack of source-content connection would result in a failure to retrieve the recalled information in the mental representation of the text.

Limitation and Further Research

One of the factors not accounted for in the current study was participants' bias towards reporting support for the topic, especially when their stances were more neutral. The prior belief rating of 4 (somewhat support) accounted for 0.36 proportion of the responses, which resulted in a highly uneven sample size per group, resulting in many non-significant differences of the sample means due to a great loss of statistical power. Future research should strive to assign participants into equal sample groups by matching their prior belief ratings to the texts that they would receive.

Another direction for future research is to investigate participants' comprehension when they read texts with both pro and con positions of the topic while listening to the agents discussing the topic. It is possible that the participants would better attend to agent sources when they experience clashes between the texts rather than clashes between compatibility of their prior beliefs to the position of the text. A condition can also be added where the sources of the texts themselves akin to the experiments done in the previous source monitoring literature are presented alongside with the agents (e.g. Braten et al., 2016; Goldman et al., 2012). If the participants could remember the sources from the text better than the agent sources, it would provide further evidence that source-content links can only be formed when sources are embedded within the text itself rather than through a secondary source of disseminated information.

Furthermore, it would be more pertinent to investigate biases of agent sources when participants are explicitly informed to attend to the agents in order to control for memory waning. If participants are told beforehand that they are to be later tested on agent sources, we could better evaluate the agent bias when the agents agree with each other that resulted in better memory for the statements. If participants unilaterally bias in having better memory for the speeches relayed by the teacher agent rather than the student in both agent agreeing and disagreeing conditions, not only will it fortify the notion of the teacher as the sole perceived source of reliable information (Rourke & Anderson, 2002), it would also bring in further questions regarding the role source memory has in stimulated cognitive disequilibrium during learning with more than one agent.

Conclusion

This dissertation investigated the impact of multiple sources of cognitive disequilibrium on learning from text with agents and using agent source monitoring as a potential mediator to resolve the cognitive disequilibrium. Unfortunately, the current result did not provide evidence that agent source monitoring mediated better comprehension during stimulated cognitive clashes between compatibility of participants' prior belief and agent disagreements. However, the current results replicated the findings from the memory literature, which collectively found that incongruences between a learner's schema and atypical information from text would result in better recognition memory. In addition, as the content of the texts is emotional (supported by high valence ratings from the participants), incongruences resulted in better recognition for surface structure. Furthermore, the current study found that prior belief congruence resulted in better summarization of the text.

The current dissertation was the first study to combine two major lines of research on text comprehension to investigate the complex dynamics between learner's prior belief, position of text, position of agents, and agent disagreements to stimulate cognitive disequilibrium for the goal of deep learning in a simulated artificial learning environment with agents. The contribution of the current study was the finding that cognitive disequilibrium stimulated by presenting agent disagreements had a positive effect only on participants' comprehension when participants were more neutral on their stance regarding topic support. An (albeit pessimistic) implication of the current results suggests that when learners have made up their minds about a particular subject or issue, exposing them to contradicting information, whether by text or by agents, can only assist them in learning the content or discriminating verbatim text; it does not seem to help the learners to have deeper comprehension where they would be able to create support for arguments and

resolve conflicts. Learners may only be able to achieve deep learning regarding information they either are unsure of their stance or have little knowledge of the subject at hand when cognitive disequilibrium, stimulated by agent disagreements, would assist in deep learning, as demonstrated from previous research (D'Mello et al., 2015). This finding is particularly relevant to the current digital informational age, with its high influx of important current controversial issues that learners would be exposed to on a regular basis-- from both social media and mainstream media news sites. Therefore, there is an urgency to find better approaches to study and formulate strategies to teach deep comprehension and critical thinking skills in order to resolve conflict and understand texts with different viewpoints. This dissertation will perhaps pave the way for future research in its kind to come in education and learning sciences.

Additional Acknowledgement

References

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review*, *7*(3), 522-530.

Anmarkrud, Ø., Braten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, *30*, 64–76.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (pp. 47-89). New York: Academic Press.

Ballard, R. (1992). Short Forms of the Marlowe-Crowne Social Desirability Scale. *Psychological Reports*, *71*(3), 1155-1160.

Begg, I. (1978). Similarity and contrast in memory for relations. *Memory & Cognition, 7*, 509-517.

Brem, A. K., Ran, K., & Pascual-Leone, A. (2013). Learning and memory. *Handbook of Clinical Neurology*, *116*, 693-737.

Benard, B. (1993). Fostering resiliency in kids. *Educational LeaderKship, 51*(3), 44-48.

Binet, A. (1900). *La suggestibilitè*. Paris: Schleicher Frères.

Bjork, E. L., & Bjork, R. A. (2014). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher and J. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society (2nd edition).* (pp. 59-68). New York: Worth.

Blanc, N., Kendeou, P., van den Broek, P., & Brouillet, D. (2008). Updating situation models during reading of news reports: Evidence from empirical data and simulation. *Discourse Processes*, *45*, 103-21.

Bloxom, M., Caul, W., Fristoe, M., & Thomson, W. (1975). On the use of student led discussion groups. *Educational Forum*, *39*, 223-230.

Braasch, J. L. G., Rouet, J., Vibert, N., & Britt, A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, *40*, 450-465.

Braten, I., Salmeron, L., Stromso, H. I. (2016). Who said that? Investigating the plausibility-induced source focusing assumption with Norwegian undergraduate readers. *Contemporary Educational Psychology*, *46*, 253-262.

Braten, I., Stromso, H. I., & Britt, M. A. (2009). Trust matters: examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, *44*, 6-28.

Britt, M. A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, *20*(4), 485–522. doi:10.1207/S1532690XCI2004_2.

Britt, M. A., & Rouet, J. F. (2012). Learning with multiple documents: Component skills and their acquisition. In M.J. Lawson and J.R. Kirby (Eds), *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes* (pp. 276-314). Cambridge University Press.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering and understanding. In J. Flavell & E. M. Markman (Eds.), *Handbook of child psychology: Cognitive development* (4th ed.; Vol. 3; pp. 515-629). New York: Wiley.

Brown, A.L., & Day, J.D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, *22*, 1-14.

Cai, Z., Feng, S., Baer, W., & Graesser, A. C. (2014). Instructional strategies in trialog-based intelligent tutoring systems. In R. Sottilare, A.C. Graesser, X. Hu, & B. Goldberg (Eds.), *Design Recommendations for Intelligent Tutoring Systems: Adaptive Instructional Strategies* (Vol.2)(pp. 225-235). Orlando, FL: Army Research Laboratory.

Cai, Z., Graesser, A.C., Forsyth, C., Burkett, C., Millis, K., Wallace, P., Halpern, D. & Butler, H. (2011). Trialog in ARIES: User Input Assessment in an Intelligent Tutoring System. In W. Chen & S. Li (Eds.), *Proceedings of the 3rd IEEE International Conference on Intelligent Computing and Intelligent Systems* (pp.429-433). Guangzhou: IEEE Press.

Cai, Z., Graesser, A.C., & Hu, X. (2015). ASAT: AutoTutor script authoring tool. In. R. Sottilare, A.C. Graesser, X. Hu, & K. Brawner (Eds.), *Design Recommendations for Intelligent Tutoring Systems: Authoring Tools* (Vol.3)(pp.199-210). Orlando, FL: Army Research Laboratory.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., Willcutt, E., & Olson, R. K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: A latent variable analysis. *Journal of Experimental Psychology General*, *141*(3), 470-88.

Cook, A. E., Halleran, J. G., & O'Brien, E. J. (1998). What is readily available during reading? A memory-based view of text processing. *Discourse Processes*, *26*(2-3), 109-29.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422-33.

Dascalu, M., Stavarache, L.L., Dessus, P., Trausan-Matu, S., McNamara, D. S., Bianco, M. (2015). Readerbench: An integrated cohesion-centered framework. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9307, 505-508.

De Pereyra, G., Britt, M. A., Braasch, J. L. G., & Rouet, J. F. (2014). Readers' memory for information sources in simple news stories: Effects of text and task features. *Journal of Cognitive Psychology, 26*, 187-204.

Deutsch, M., Coleman, P. T., & Marcus, E. C. (2006). *The handbook of conflict resolution: Theory and practice*. San Francisco, CA: John Wiley & Sons, Inc.

D'Mello, S. K., Dale, R., & Graesser, A. C. (2012). Disequilibrium in the mind, disharmony in the body. *Cognition and Emotion, 26*(2), 362-374.

D'Mello, S. K., & Graesser, A. C. (2012). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems, 2*(4), 1-38.

D'Mello, S. K., & Graesser, A. C. (2014). Confusion and its dynamics during device comprehension with breakdown scenarios. *Acta Psychologica*, *151*, 106-116. doi:10.1016/j.actpsy.2014.06.005

D'Mello, S. K., Lehman, B., Pekrun, R., & Graesser, A. C. (2015). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153-170.

D'Mello, S. K., Lehman, B., Pekrun, R., & Graesser, A. C. (2015). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153-170.

Epstein, M.L., Phillips, W.D., & Johnson, S.J. (1975). Recall of related and unrelated words as a function of processing level. *Journal of Experimental Psychology: Human Learning and Memory, 1*, 149-152.

Feng, S., Cai, Z., & Graesser, A. C. (2017). Detecting contradiction in agent source monitoring during expository text comprehension. *Paper presented at the Annual Meeting of the Society for Computers in Psychology*. Vancouver: British Columbia.

Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2007). Modeling multiple levels of text presentation. In F. Schmalhofer, & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 133-157). Mahwah, NJ: Erlbaum.

Fulmer, S. M., D'Mello, S. K., Stratin, A., & Graesser, A. C. (2015). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary Educational Psychology*, *41*, 98-110.

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehension and learning from internet sources: Processing patterns of better and poorer learners. *Reading Research Quartely*, *47*(4), 356-381.

Goldman, S. R., Ozuru, Y., Braasch, J. L. G., Manning, F., Lawless, L., Gomez, K. et al. (2011). Literacies for learning: A multiple source comprehension illustration. In N. L. S. Stein & S. W. Raudenbush (Eds.), *Developmental science goes to school: Implications for policy and practice* (pp. 30-44). New York, NY: Routledge.

Goldstein, H. (2011). *Multilevel Statistical Models*. West Sussex, UK: John Wiley & Sons Ltd.

Graesser, A.C. (2015). Deeper learning with advances in discourse science and technology. *Policy Insights from Behavioral and Brain Sciences, 2*, 42-50.

Graesser, A.C. (2016). Conversations with AutoTutor help students learn. *International Journal of Artificial Intelligence in Education, 26,* 124-132.

Graesser, A.C., Baer, W., Feng, S., Walker, B., Clewley, D., Hays, D.P., & Greenberg, D. (2015). Emotions in Adaptive Computer Technologies for Adults Improving Reading. In S. Tettegah and M. Gartmeier (Eds.), *Emotions, Technology, Design, and Learning* (pp. 3-25). New York: Elsevier.

Graesser, A.C., Cai, Z., Baer, W., Olney, A.M., Hu, X., Reed, M., & Greenberg, D. (2016). Reading Comprehension Lessons in AutoTutor for the Center for the Study of Adult Literacy. In S. Crossley and D. McNamara (Eds.), *Adaptive Educational Technologies for Literacy Instruction* (pp. xx-yy). New York: Routledge.

Graesser, A. C., Chipman, P., Leeming, F., & Biedenbach, S. (2009). Deep learning and emotion in serious games. In U. Ritterfeld, M. Cody, & P. Vorderer (Eds.), *Serious games: Mechanisms and effects* (pp. 81-100). New York and London: Routledge, Taylor & Francis.

Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Recognition memory for typical and

    atypical actions in scripted activities: Tests of a script pointer plus tag hypothesis.

    Journal of verbal learning and verbal behavior, *18*, 319-332.

Graesser, A. C., Li, H., Feng, S. (2015). Constructing inferences in naturalistic reading contexts.

    In E. O'Brien, A. Cook, and R. Lorch, (Eds.), *Inferences during Reading* (pp. 290-320).

    Cambridge: Cambridge University Press.

Graesser, A. C., Li, H., & Forsyth, C. (2014). Learning by communicating in natural language

    with conversational agents. *Current Directions in Psychological Science*, *23*(5), 374-380.

    doi: 10.1177/0963721414540680.

Graesser, A. C., & Mandler, G. (1975). Recognition memory for the meaning and surface

    structure of sentences. *Journal of Experimental Psychology: Human Learning and*

    Memory, *1*(3), 238-248.

Graesser, A. C., & Nakamura, G. V. (1982). The impact of a schema on comprehension and

    memory. *Psychology of Learning and Motivation*, *16*, 59-109.

Graesser, A., Ozuru, Y., & Sullins, J. (2010). What is a good question? In M. McKeown & G.

    Kucan (Eds.), *Bringing reading research to life* (pp. 112-141). New York: Guilford.

Graesser, A. C., Singer, M., & Trabasso, T (1994). Constructing inferences during narrative text

    comprehension. *Psychological Review*, *101*, 371-95.

Graesser, A. C., Woll, S. B., Kowalski, D. J., & Smith, D. A. (1980). Memory for typical and

    atypical actions in scripted activities. *Journal of Experimental Psychology: Human*

    Learning and Memory, *6*, 503–515.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness.

 Journal of Memory and Language, 32, 421-445.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis.

 Psychological Methods, *15*(4), 309.

Jeong, A. C. (2003). The sequential analysis of group interaction and critical thinking in online

 threaded discussions. *The American Journal of Distance Education*, *17*(1), 25-43.

Johnson, D. W., & Johnson R. T. (2002). Social interdependence theory and university

 instruction - Theory into practice. *Swiss Journal of Psychology*, *61*(3), 119-129.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source Monitoring. *Psychological*

 Bulletin, *114*(3), 3-28.

Kiili, C., Laurinen, L., & Marttunen, M. (2008). Students evaluating Internet sources: From

 versatile evaluators to uncritical readers. *Journal of Educational Computing Research*,

 39(1), 75–95. doi:10.2190/EC.39.1.e

Kendeou, P., Braasch, J. L. G., & Bråten, I. (2016). Optimizing conditions for learning: Situating

 refutations in epistemic cognition. *Journal of Experimental Education*, *84*, 245–263.

Kendeou, P., & O'Brien, E. J. (2014). The knowledge revision components (KReC) framework:

 Processes and mechanisms. In D. N. Rapp & J. L. G. Braasch (Eds.), *Processing*

 inaccurate information: Theoretical and applied perspectives from cognitive science and

 the educational sciences (pp. 353–377). Cambridge, MA: MIT Press.

Kendeou, P., Walsh, E. K., Smith, E. R., & O'Brien, E. J. (2014). Knowledge revision processes

 in refutation texts. *Discourse Processes*, *51*, 374–397.

Kensinger, E. A. (2007). How negative emotion affects memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science*, *16*, 213–218.

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, *53*, 983-997.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.

Kremer, J. & McGuinness, C. (1998). Cutting the Cord: Student-Led Discussion Groups in Higher Education. *Education and Training*, *40*(2), 44-49.

Lawless, K.A., Goldman, S.R., Gomez, K., Manning, F., & Braasch, J. (2012). Assessing multiple source comprehension through evidence centered design. In J.P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Reaching an understanding: Innovations in how we view reading assessment* (pp. 3-17). Lanham, MD: R&L Education.

Lehman, B. A., D'Mello, S. K., Strain, A., Gross, M., Dobbins, A., Wallace, P., Millis, K., & Graesser, A. C. (2011). Inducing and tracking confusion with contradictions during critical thinking and scientific reasoning. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Proceedings of 15th International Conference on Artificial Intelligence in Education* (pp. 171-178). Berlin: Springer-Verlag.

Lehman, B., & Graesser, A. C. (2014). Impact of agent role on confusion induction and learning. In S. Trausan-Matu, K. Boyer, M. Crosby, & K. Panou (Eds.), *Proceedings of the 12th International Conference on intelligent tutoring systems* (ITS 2014) (pp. 45 - 54). Berlin: Springer.

Long, D. L. (1989). *Recognition memory for conversations embedded in narrative* (Doctoral dissertation, University of Memphis). Retrieved from https://www.researchgate.net/publication/36060090_Recognition_memory_for_conversations_embedded_in_narrative

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of educational psychology, 106(4), 901-918. doi: http://dx.doi.org/10.1037/a0037123

Magoon, A. J. (1977). Constructivist approaches in educational research. *Review of Educational Research*, *47*(4), 651-93.

Maier, J., & Richter, T. (2013). Text belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction*, *31*, 151-175.

Matsa, K. E., & Lu, K. (2016, September 14). 10 facts about the changing digital news landscape. *Pew Research Center*. Retrieved from http://www.pewresearch.org/fact-tank/2016/09/14/facts-about-the-changing-digital-news-landscape/

Meade, M. L., & Roediger, H. L. (2002). Explorations in the social contagion of memory. *Memory & Cognition*, *30*(7), 995-1009.

National Research Council (2011). *Improving Adult Literacy Instruction: Options for Practice and Research*. Washington, D.C.: The National Academies Press.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1200-1210.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer Verlag.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodating of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211-227.

Rapp, D. N. (2008). How do readers handle incorrect information during reading? *Memory & Cognition*, *36*, 688-701.

Riefer, D. M., & Batchelder, W. H. (1995). A multinomial modeling analysis of the recognition-failure paradigm. *Memory & Cognition*, *23*(5), 611-630.

Roediger, H. L., Meade, M. L., & Bergman, E. T. (2001). The social contagion of memory. *Psychonomic Bulletin & Review*, *8*, 365-371.

Rouet, J. (2006). *The skills of document use: From text comprehension to Web-based learning.* Mahwah, NJ: Erlbaum.

Rourke, L., & Anderson, T. (2002). Using peer teams to lead online discussions. *Journal of Interactive Media in Education*, *1*, Retrieved on August 31, 2011 from http://jime.open.ac.uk/article/2002-1/80.

Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, *2*, 437-442.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory and Cognition*, *2*, 95-100.

Schank, R. C., & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, New Jersey: Erlbaum.

Schermerhorn, S. M., Goldschmid, M. L., & Shore, B. M. (1976). Peer Teaching in the Classroom: Rationale and Feasibility. *Improving Human Performance Quarterly*, *5*, (1), 27-34.

Smith, D. A., & Graesser, A. C. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. *Memory and Cognition*, *9*, 550–559.

Stadtler, M. & Scharrer, L., Brummernhenrich, B., & Bromme, R. (2013). Dealing with uncertainty: Readers' memory for and use of conflicting information from science texts as function of presentation format and source expertise. *Cognition and Instruction, 31*(2), 130-150.

Stein, N. L., & Miller, C. A. (1990) 'I win - you lose: The development of argumentative thinking. In J. F. Voss, D. N. Perkins, & J. Segal (Eds.), *Informal Reasoning and Instruction* (pp.265-290). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review*, *13*(6), 1023-1027.

Strømsø, H.I., & Bråten, I., & Britt, M.A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction, 18*, 513-527.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis.

Voss, J. F., & Wiley, J. (2000). A case study of developing historical understanding via instruction: The importance of integrating text components and constructing arguments. In P. Stearns, S. Wineburg, and P. Seixas (Eds.), *Knowing, teaching and learning in History* (pp. 375-389). New York: NYU Press.

Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, *91*(2), 301-311.

Wiley, J., Goldman, S., Graesser, A., Sanchez. C., Ash, I., & Hemmerich, J. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal, 46*, 1060-1106.

Example Compatible Text and the Corresponding Agent Speech Pairs.

Topic: Black Lives Matter Movement

Condition: text Compatible teacher to text compatible student to text incompatible (text pro T+S-)

Page 1:

Black humanity and dignity requires Black political will and power. Despite constant exploitation and perpetual oppression, Black people have bravely and brilliantly been the driving force pushing the U.S. towards the ideals it articulates but has never achieved. In recent years we have taken to the streets, launched massive campaigns, and impacted elections, but our elected leaders have failed to address the legitimate demands of our Movement. We can no longer wait.

In response to the sustained and increasingly visible violence against Black communities in the U.S. and globally, a collective of more than 50 organizations representing thousands of Black people from across the country have come together with renewed energy and purpose to articulate a common vision and agenda. We are a collective that centers and is rooted in Black communities, but we recognize we have a shared struggle with all oppressed people; collective liberation will be a product of all of our work.

Teacher (+): The black community needs a movement that will address the issues of systemic oppression and violence against them.

Student (-): The black community needs to prove that oppression has committed against them before claiming legitimacy of a movement.

Page 2:

We believe in elevating the experiences and leadership of the most marginalized Black people, including but not limited to those who are women, queer, trans, femmes, gender nonconforming, Muslim, formerly and currently incarcerated, cash poor and working class, differently-abled, undocumented, and immigrant. We are intentional about amplifying the particular experience of state and gendered violence that Black queer, trans, gender nonconforming, women and intersex people face. There can be no liberation for all Black people if we do not center and fight for those who have been marginalized. It is our hope that by working together to create and amplify a shared agenda, we can continue to move towards a world in which the full humanity and dignity of all people is recognized.

While this platform is focused on domestic policies, we know that patriarchy, exploitative capitalism, militarism, and white supremacy know no borders. We stand in solidarity with our

international family against the ravages of global capitalism and anti-Black racism, human-made climate change, war, and exploitation.

Teacher (+): The elevated leadership of the black community will work with other marginalized people, which will be good for all.

Student (-): The increasing leadership of the black community defines who they mean by marginalized people, which is not good.

Page 3:

We also stand with descendants of African people all over the world in an ongoing call and struggle for reparations for the historic and continuing harms of colonialism and slavery. We also recognize and honor the rights and struggle of our Indigenous family for land and self-determination.

We have created this platform to articulate and support the ambitions and work of Black people. We also seek to intervene in the current political climate and assert a clear vision, particularly for those who claim to be our allies, of the world we want them to help us create. We reject false solutions and believe we can achieve a complete transformation of the current systems, which place profit over people and make it impossible for many of us to breathe.

Together, we demand an end to the wars against Black people. We demand that the government repair the harms that have been done to Black communities in the form of reparations and targeted long-term investments. We also demand a defunding of the systems and institutions that criminalize and cage us. This document articulates our vision of a fundamentally different world.

Teacher (+): The movement seeks to change the world into a better system where black people will stop being criminalized by the institution that cages them.

Student (-): It seems that the movement wishes to defund the criminal justice system, which will let criminals roam free on the streets, making it unsafe for all.

Page 4.

However, we recognize the need to include policies that address the immediate suffering of Black people. These policies, while less transformational, are necessary to address the current material conditions of our people and will better equip us to win the world we demand and deserve.

We recognize that not all of our collective needs and visions can be translated into policy, but we understand that policy change is one of many tactics necessary to move us towards the world we

94

envision. We have come together now because we believe it is time to forge a new covenant. We are dreamers and doers and this platform is meant to articulate some of our vision. The policy briefs we have drafted also elevate the brave and transformative work our people are already engaged in, and build on some of the best thinking in our history of struggle. This agenda continues the legacy of our ancestors who pushed for reparations, Black self-determination and community control; and also propels new iterations of movements such as efforts for reproductive justice, holistic healing and reconciliation, and ending violence against Black cis, queer, and trans people.

Teacher (+):The movement seeks to change policy that will assist black people in getting the justice that they deserve

Student (-): Policies need to be changed to ensure justice for all races, not just black people.

Condition: text Incompatible teacher to text compatible student to text incompatible (text Incompatible T+S-)

Page 1

A large group of Black Lives Matter organizations have come together to form the Movement For Black Lives. This coalition released their new policy agenda with six core demands and forty policy priorities. Movement For Black Lives has put together a, "clear vision of the world where black humanity and dignity is the reality."

The list of their demands reads like the Christmas wish list of a kid who just went through a giant toy catalog, "I'll take that, and that, and that." Rather than actually writing a proposal that can be taken seriously, they came up with a plan that, if executed, would ensure that no black person would ever need to work or be held accountable for any crimes ever again.

The six demands are:

1.   End the war on black people.

2.   Reparations for past and continuing harms.

3.   Divestment from the institutions that criminalize, cage, and harm black people; and investment in the educations, health and safety of black people.

4.   Economic justice for all and reconstruction of the economy to ensure our communities have collective ownership, not merely access.

5.   Community control of the laws, institutions, and policies that most impact us.

6.   Independent black political power and black self-determination in all areas of society.

Teacher (+): The list of demands from the movement is unrealistic and denotes a complete disinterest in negotiating.

Student (-): The list of demands from the movement addresses pressing problems and therefore there is no time for negotiating.

Page 2:

So let's break this down. The first demand is an end to the war on black people. The Movement For Black Lives page outlines this demand as a demand for an "end to the criminalization, incarceration, and killing of our people" which includes:

• An end to the criminalization and dehumanization of black youth.

• End to zero-tolerance school policies and arrests of students

• The removal of police from schools

• The relocation of funds from police and punitive school discipline practices to restorative services.

• An end to capital punishment

• An end to money bail, mandatory fines, fees, court surcharges and defendant-funded court proceedings.

• An end to the use of past criminal history to determine eligibility for housing, education, licenses, voting, loans, employment, and other services and needs.

• The demilitarization of law enforcement including law enforcement in schools and on college campuses.

• An end to the privatization of police, prisons, jails, probation, parole, food, phone and all other criminal justice related services.

• An end to public jails, detention centers, youth facilities, and prisons.

From reading their list, the "war on black people" appears to be also known as, "holding people accountable for their actions." They don't want to be bothered while committing crime. Then if they are arrested for a crime, they want the law-abiding citizens to pay for it, and then if they are sentenced, they don't want to be fined or jailed for their crimes.

Teacher (+): What's included in first demand from the movement is essentially a severe reduction of law enforcement and criminal justice system.

Student (-): The first demand from the movement only includes that are considered to be basic human rights.

96

Page 3:

The second demand is for "reparations for past and continuing harms." This basically means that, they want white people to pay every black person tens of thousands of dollars each year in cash, free land, free food, and free education just for being black. And no, that's not an over-exaggeration, that's exactly what they asked for.

The third demand is for divestment from the institutions that criminalize, cage, and harm black people; and investment in the educations, health and safety of black people. This demand gets even crazier. It calls for a reallocation of funds at the federal, state and local level from policing and incarceration. The decriminalization, immediate release, and record expungement of all drug related offenses and prostitution. This alone could release thousands of violent drug dealers instantly back on the street. We better get the reparations in place first to prevent all the dope-fiends from robbing people to get their drug fix.

The fourth demand is for "economic justice for all and reconstruction of the economy to ensure our communities have collective ownership, not merely access." This demand calls for a progressive restructuring of tax codes to ensure a radical and sustainable redistribution of wealth. This part of the plan basically reads like a socialist's ideal political structure (real socialists, not that Bernie Sanders brand of socialism.) The big exception though, is that it calls for segregating businesses by the race. And then, all aid in the form of grants, loans or contracts to help facilitate this plan must go to black led or black supported networks and organizations as defined by communities. So if you own a business and a black person doesn't approve of your business you won't be eligible for this plan.

Teacher (+): The movement demands reparations based on race. Specifically, whites need to pay money to blacks.

Student (-): The article exaggerates what entails the reparation that black people deserves from being harmed by the system.

Page 4:

The fifth demand is for "community control of the laws, institutions, and policies that most impact us," and, "community control of local, state, and federal law enforcement agencies." Ensuring that communities most "harmed" by "destructive" policing have the power to hire and fire officers, determine disciplinary action, control budgets and policies, and subpoena relevant agency information. This one is the most cringe worthy. The ability to hire and fire officers and determine disciplinary action. We've seen time and time again the power of the machine to condemn an officer for their actions before the facts of the case were release. And that doesn't even consider that the facts don't matter to these people, even after the truth comes forward (Michael Brown.) This policy would effectively eliminate law enforcement all together, which Black Lives Matter has repeatedly stated is their actual goal.

The sixth and final demand is for independent black political power and black self-determination in all areas of society. This demand states a vision of remaking the current U.S. political system in order to create a real democracy where black people can effectively exercise full political power. This ignored the fact that President is black, the US Attorney General is black, the black Supreme Court justice, not to mention senators and representatives. It seems that they effectively want more political control of the country than any other ethnicity has. The demand goes on to call for an end to the criminalization of black political activity and immediate release of all political prisoners. That basically means that they want to be able to riot and block roadways without consequences.

This is Black Lives Matter's vision of the future, where criminals are the victims. If you have dark skin, then the rest of the world owes you something. Where the laws are not enforced against certain people, based on the color of their skin. When the general public hears "Black Lives Matter," they often think, "Of course the lives of black people matter, I support that." These people are generally unaware of what Black Lives Matter groups' policy agendas, or that these groups appear to be getting more extreme by the day.

Please spread the word that this is a group that should be shunned as extremists, not supported.

Teacher (+): The movement's real goal is to have black communities create and regulate criminal laws independent of the state.

Student (-): The laws from the state are flawed, therefore it is only right that the communities should have autonomy of their own affairs.

APPENDIX B

Example Old-New Item Counterbalancing Scheme for Source Monitoring Recognition Test.

Subject	Condition	Old Item	New Item
1	AproT+S-	T: The black community needs a movement that will address the issues of systemic oppression and violence against them. S: The black community needs to prove that oppression has committed against them before claiming legitimacy of a movement.	T: The oppression that black people face is deeply rooted in the constant exploitation of black communities across the U.S. S: Black people need to prove that they are facing constant exploitation in their communities within the U.S.
1	AproT+S-	S: The increasing leadership of the black community defines who they mean by marginalized people, which is not good. T: The elevated leadership of the black community will work with other marginalized people, which will be good for all.	S: Those in the black community who gets to define who is marginalized and who isn't has an unjustified power. T: The elevated leadership of the black community hopes to work together to ensure humanity and dignity of all people.
1	AproT+S-	T: The movement seeks to change the world into a better system where black people will stop being criminalized by the institution that cages them. S: It seems that the movement wishes to defund the criminal justice system, which will let criminals roam free on the streets, making it unsafe for all.	T: The black community deserves reparations for the historic colonialism and slavery that has been done to the black people. S: People do not deserve reparations for the harms that has been done to their ancestors and not directly to them.

1	AproT+S-	S: Policies need to be changed to ensure justice for all races, not just black people. T: The movement seeks to change policy that will assist black people in getting the justice that they deserve.	S: The movement uses identity politics to separate people based on race, which will result in more conflict that they will then thrive on. T: The movement seeks to build self-determination and community control for the black people, which will advance the society as a whole.
2	AproT+S-	T: The oppression that black people face is deeply rooted in the constant exploitation of black communities across the U.S. S: Black people need to prove that they are facing constant exploitation in their communities within the U.S.	T: The black community needs a movement that will address the issues of systemic oppression and violence against them. S: The black community needs to prove that oppression has committed against them before claiming legitimacy of a movement.
2	AproT+S-	S: Those in the black community who get to define who is marginalized and who isn't have unjustified power. T: The elevated leadership of the black community hopes to work together to ensure humanity and dignity of all people.	S: The increasing leadership of the black community defines who they mean by marginalized people, which is not good. T: The elevated leadership of the black community will work with other marginalized people, which will be good for all.
2	AproT+S-	T: The black community deserves reparations for the historic colonialism and slavery that has been done to the black people. S: People do not deserve reparations for the harms that has been done to their ancestors and not directly to them.	T: The movement seeks to change the world into a better system where black people will stop being criminalized by the institution that cages them. S: It seems that the movement wishes to defund the criminal justice system, which will let criminals roam

			free on the streets, making it unsafe for all.
2	AproT+S-	S: The movement uses identity politics to separate people based on race, which will result in more conflict that they will then thrive on. T: The movement seeks to build self-determination and community control for the black people, which will advance the society as a whole.	S: Policies need to be changed to ensure justice for all races, not just black people. T: The movement seeks to change policy that will assist black people in getting the justice that they deserve.

Example Sentence Recognition Items for the Compatible Text on the Black Lives Matter Movement (order of multiple-choice items will be randomized).

Which of the following sentences appeared in one of the texts that you've read (Question to the Participant)?

1. **Please choose from the following:**
   A. In response to the sustained and increasingly visible violence against Black communities in the U.S. and globally, a collective of more than 50 organizations representing thousands of Black people from across the country have come together with renewed energy and purpose to articulate a common vision and agenda. (S+M+)
   B. A collection of more than 50 organizations representing thousands of Black people from across the country, in response to the sustained and increasingly visible violence against Black communities in the U.S. and globally, have come together with reviewed energy and purpose to articulate a common vision and agenda. (S-M+)
   C. In response to the sustained and increasingly visible violence against Black communities in the U.S. and globally, a collective of more than 50 organizations representing thousands of Black people from across the country have come together with renewed energy and purpose to express individual people's vision and agendas. (S+M-)
   D. A collection of more than 50 organizations representing thousands of Black people from across the country, in response to the sustained and increasingly visible violence against Black communities in the U.S. and globally, have come together with reviewed energy and purpose to express individual people's vision and agendas. (S-M-)

2. **Please choose from the following:**

   A. We are intentional about amplifying the particular experience of state and gendered violence that Black queer, trans, gender nonconforming, women and intersex people face. (S+M+)
   B. Black queer, trans, gender nonconforming, women and intersex people face a particular experience of state and gendered violence, that we are intentional about amplifying. (S-M+)
   C. We are intentional about amplifying all individual experiences of state and gendered violence including those faced by Black queer, trans, gender nonconforming, women and intersex people. (S+M-)
   D. Black queer, trans, gender nonconforming, women and intersex people are among the target of the state and gendered violence, and we are intentional about amplifying every experience that individuals faced. (S-M-)

3. Please choose from the following:

A. We reject false solutions and believe we can achieve a complete transformation of the current systems, which place profit over people and make it impossible for many of us to breathe. (S+M+)
B. We believe we can achieve a complete transformation of the current systems, which place profit over people and make it impossible for many of us to breathe, and we reject all false solutions. (S-M+)
C. We reject false solutions and we are aware that we cannot make a complete transformation of the current systems, which place profit over people and make it impossible for many of us to breathe. (S+M-)
D. We are aware that we cannot achieve a complete transformation of the current systems, which place profit over people and make it impossible for many of us to breathe, but we reject all false solutions. (S-M-)

4. Please choose from the following:

A. The policy briefs we have drafted also elevate the brave and transformative work our people are already engaged in, and build on some of the best thinking in our history of struggle. (S+M+)
B. Our people are already engaged in the brave and transformative work, and we have drafted the policy briefs to also elevate and build on some of the best thinking in our history of struggle. (S-M+)
C. The policy briefs we have drafted also inspire the brave and transformative work our people should start engaging in, and build on some of the best thinking in our history of struggle. (S+M-)
D. Our people are also inspired from and should start engaging in the brave and transformative work, and we have drafted the policy briefs to build on some of the best thinking in our history of struggle. (S-M-)

APPENDIX D

Full Counterbalancing Scheme of the Current Study Design.

Participant	Topic	Text Position	Agent Speech Condition	Teacher Agent
1	A	Pro	T+S-	Cristina
1	B	Con	T-S+	Cristina
1	C	Pro	T+S+	Cristina
1	D	Con	T-S-	Cristina
2	A	Con	T+S-	Cristina
2	B	Pro	T-S+	Cristina
2	C	Con	T+S+	Cristina
2	D	Pro	T-S-	Cristina
3	A	Pro	T-S+	Cristina
3	B	Con	T+S-	Cristina
3	C	Pro	T-S-	Cristina
3	D	Con	T+S+	Cristina
4	A	Con	T-S+	Cristina
4	B	Pro	T+S-	Cristina
4	C	Con	T-S-	Cristina
4	D	Pro	T+S+	Cristina
5	B	Pro	T+S+	Cristina
5	C	Con	T-S+	Cristina
5	D	Pro	T+S-	Cristina
5	A	Con	T-S-	Cristina
6	B	Con	T+S+	Cristina
6	C	Pro	T-S+	Cristina
6	D	Con	T+S-	Cristina

6	A	Pro	T-S-	Cristina
7	B	Pro	T-S-	Cristina
7	C	Con	T+S-	Cristina
7	D	Pro	T-S+	Cristina
7	A	Con	T+S+	Cristina
8	B	Con	T-S-	Cristina
8	C	Pro	T+S-	Cristina
8	D	Con	T-S+	Cristina
8	A	Pro	T+S+	Cristina
9	C	Pro	T+S+	Jordan
9	D	Con	T-S-	Jordan
9	A	Pro	T+S-	Jordan
9	B	Con	T-S+	Jordan
10	C	Con	T+S+	Jordan
10	D	Pro	T-S-	Jordan
10	A	Con	T+S-	Jordan
10	B	Pro	T-S+	Jordan
11	C	Pro	T-S-	Jordan
11	D	Con	T+S+	Jordan
11	A	Pro	T-S+	Jordan
11	B	Con	T+S-	Jordan
12	C	Con	T-S-	Jordan
12	D	Pro	T+S+	Jordan
12	A	Con	T-S+	Jordan
12	B	Pro	T+S-	Jordan
13	D	Pro	T+S-	Jordan
13	A	Con	T-S-	Jordan

13	B	Pro	T+S+	Jordan
13	C	Con	T-S+	Jordan
14	D	Con	T+S-	Jordan
14	A	Pro	T-S-	Jordan
14	B	Con	T+S+	Jordan
14	C	Pro	T-S+	Jordan
15	D	Pro	T-S+	Jordan
15	A	Con	T+S+	Jordan
15	B	Pro	T-S-	Jordan
15	C	Con	T+S-	Jordan
16	D	Con	T-S+	Jordan
16	A	Pro	T+S+	Jordan
16	B	Con	T-S-	Jordan
16	C	Pro	T+S-	Jordan

THE UNIVERSITY OF
MEMPHIS™

Institutional Review Board
Office of Sponsored Programs
University of Memphis
315 Admin Bldg
Memphis, TN 38152-3370

PI: Arthur Graesser
Co-Investigator:
Advisor and/or Co-PI:
Department: Psychology
Study Title: Understanding the cognitive and motivational profiles of struggling adult readers and developing effective and engaging literacy programs to address their literacy learning needs
IRB ID: 2255
Submission Type: Renewal
Level of Review: Expedited

IRB Meeting Date:
Decision: Approved
Approval Date: Jul 7, 2017
Expiration Date: Jul 7, 2018

Research Notes:
Findings:

The IRB has reviewed the renewal request.

Approval of this project is given with the following obligations:
1. If this IRB approval has an expiration date, an approved renewal must be in effect to continue the project prior to that date. If approval is not obtained, the human consent form(s) and recruiting material(s) are no longer valid and any research activities involving human subjects must stop.
2. When the project is finished or terminated, a completion form must be completed and sent to the board.
3. No change may be made in the approved protocol without prior board approval, whether the approved protocol was reviewed at the Exempt, Expedited or Full Board level.
4. Exempt approval are considered to have no expiration date and no further review is necessary unless the protocol needs modification.

Thank you,
James P. Whelan, Ph.D.

Institutional Review Board Chair
The University of Memphis.
Note: Review outcomes will be communicated to the email address on file. This email should be considered an official communication from the UM IRB.