2020

Examining Boundary Conditions to the attitude Consistency Effect

Dylan Blaum
dylblaum@gmail.com

ABSTRACT

EXAMINING BOUNDARY CONDITIONS TO THE ATTITUDE CONSISTENCY EFFECT

Dylan Blaum, Ph.D.
Department of Psychology
Northern Illinois University, 2020
Anne Britt, Director

Research on the impact of attitudes on argument processing has found that attitudes and beliefs can impact evaluation and processing, but has not deeply explored boundary conditions to this *attitude consistency effect*. This dissertation investigates three types of boundary conditions to this effect: whether the quality of the argument matters (argument quality), whether the individual's strength of that attitude matters (attitude strength), and whether the evaluative nature of the task matters (task).

In Experiment 1, I found some support for argument quality as a boundary condition to the attitude consistency effect. Attitude consistent arguments were rated as higher quality/stronger than attitude inconsistent arguments, but only for arguments of high quality (warranted). Additionally, some support for attitude strength was found in Experiment 1. Arguments that were about topics that participants weakly or strongly agreed with showed higher ratings when attitude consistent compared to attitude inconsistent. For arguments about topics which participants weakly or strongly disagreed, there was no attitude consistency effect.

In Experiment 2, I found support for task and attitude strength as boundary conditions to the attitude consistency effect, but did not completely replicate the interaction between argument quality and attitude consistency on evaluations found in Experiment 1. Being in an evaluative task lead to higher reading time for argument reasons compared to a control memory task. While I did not replicate the interaction between argument quality and attitude consistency on

evaluations, a three-way interaction between attitude strength, quality, and attitude consistency on reason reading time showed an attitude consistency effect for weakly disagree, weakly agree, and strongly agree warranted arguments.

These results are important when trying to teach/train people to evaluate arguments because people may need to be taught to manage their attitudes in order to prevent them from interfering during a logical evaluation task. Additionally, these results also provide some support for existent models related to attitudes and argument processing.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

AUGUST 2020

EXAMINING BOUNDARY CONDITIONS TO THE ATTITUDE CONSISTENCY EFFECT

BY

DYLAN BLAUM

A DISSERATION SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

Doctoral Director:

Anne Britt

ACKNOWLEDGEMENTS

There are countless people who helped me succeed in this project. First and foremost I have to thank my advisor, Dr. Anne Britt. Without her guidance, push, and understanding I would not have been able to get to this point in my program to even think about completing this dissertation. I would also like to thank the rest of my dissertation committee: Dr. Keith Millis, Dr. Katja Wiemer, and Dr. Brad Sagarin. I am extremely grateful for their feedback as well as their flexibility and understanding when unforeseen circumstances led to last minute changes in the project. Of course, I would also like to show thanks to my lab mates Katy and Christian. Being able to bounce ideas off of them and think through tough problems helped to keep the project moving. I would like to thank my friend Zak for giving me an outlet to clear my mind (and clear shuttles) and take a step away from school to play and coach badminton. Finally, I would like to thank my friend Alie for all of the emotional and mental support she gave me to help keep me on track.

DEDICATION

This dissertation is dedicated to my nieces, Cadence and Norah, as well as my nephew Mason.

I’m Doctor Uncle Dylan now!

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER ONE: INTRODUCTION & REVIEW OF THE LITERATURE

The internet, especially social media, has become a ubiquitous part of our everyday lives. These tools allow us to share many things like videos, music, art, and memes with friends, family, and strangers online. Another thing that people enjoy sharing online is their beliefs and attitudes about current hot topics. This results in people being exposed to arguments about topics ranging from, whether or not vaccinations should be mandatory to whether or not recycling laws should be updated. When people read these arguments they will often engage with those who posted, either to discuss their agreement or disagreement with the argument. While engaging with these posts, a number of things may influence how people respond. Britt and Larson (2003) suggested that there are some features that may prompt readers to identify that they are reading a claim and thus process what they are reading as an argument. These features include modals, qualifiers, attitude consistency, controversiality, and presence of a reason to support the claim. There has been little investigation into how these features may influence argument processing. One of the features that has been explored is the role that beliefs, values, and attitudes play in argument processing.

Some research has investigated how people's beliefs about topics influence how they process and evaluate arguments about that topic (Clark et al., 2008; Edwards & Smith, 1996; Wolfe & Kurby, 2017). When reading arguments that one disagrees with, people will rate the arguments as weaker, spend more time reading, and produce more counter arguments compared to arguments they agree with (Edwards & Smith, 1996). According to Edwards and Smith (1996), this *disconfirmation bias* occurs as a result of people scrutinizing the arguments when

reading. This scrutiny involves strategically processing arguments to find flaws in order to discount the argument that one disagrees with and preserve one's belief on the other side of the argument. Clark and Wegener (2013) call this processing *defending* but also suggest that under some circumstances people will engage in effortful processing, not to defend their pre-existing strong beliefs, but to *bolster*, or strengthen, their weak beliefs. Both defending (scrutiny) and bolstering require the reader to engage in some form of effortful processing (Hasher & Zacks, 1979). This dissertation was designed to explore three boundary conditions to these effects of prior attitudes on argument processing and evaluation. Research question 1 addresses a potential boundary condition based on the type of materials. It is important to understand whether attitudes would even affect the acceptance of arguments with irrelevant reasons. It may be that arguments have to reach a threshold of acceptability to have attitudes have an effect or it may be that attitudes motivate the reader to make an irrelevant reason seem better by way of extended inferencing. Research question 2 addresses a boundary based on an individual's attitude strength. Perhaps attitudes are more influential when individuals have more strongly held attitudes about the topics of the argument compared to when they are more ambivalent or hold weaker attitudes. Research question 3 addresses whether attitudes influence argument processing when the task requires comprehension and not evaluation. These answers can help us understand how people process arguments and when their attitudes about the topics impact that processing.

Argument Structure

Before digging too far into the weeds, it's important to understand what I mean by arguments. For the purpose of this dissertation I will be referring to informal arguments that take a structure similar to that laid out by Toulmin (1958). Generally, arguments in this structure take a simple form where a claim is supported by a reason. The reason is able to provide support for

the claim because of an underlying, and often unstated, *warrant*. The warrant acts to connect the reason to the claim in a meaningful and supportive way. This should become clearer in the examples below.

These three components work together to form the framework structure for informal arguments. For example, consider arguments 1a and 1b with the policy claims about recycling.

1a. Recycling should be federally mandated because less material will go into our landfills.

Argument 1a can be dissected as follows:

Claim: Recycling should be federally mandated

Reason: because less material will go into our landfills.

Unstated Warrant: *Actions that reduce the amount of material that goes into our landfills should be federally mandated.*

1b. Recycling should not be federally mandated because different types of materials can be recycled.

Argument 1b can be dissected as follows:

Claim: Recycling should not be federally mandated

Reason: because different types of materials can be recycled.

Unstated Warrant: *Actions that can be done with different types of materials should not be federally mandated.*

The claim of Argument 1a is that 'recycling should be federally mandated.' This claim is composed of two major pieces, the *theme* and the *predicate*. The theme of an argument is simply the topic of the argument. In the case of Argument 1a and 1b, the theme of the argument is recycling since both arguments are about recycling. The predicate of an argument is the main

verb phrase of the claim, in the case of Argument 1a, “should be federally mandated.” The predicate states the action that acts on the theme in a claim. In the case of both arguments above, the predicate includes a “should” in its verb phrase. This ‘should’ is a marker that typically indicates the claim being made is a *policy* claim. There are other claim types found in informal argumentation, such as factual and evaluative, and they have different features and markers. For this dissertation I am going to exclusively focus on policy claims since they are commonly used in research and are also very common within the current political discourse. The predicate does not only indicate the type of claim being made, it also indicates the *side* that the author is taking, whether the author is for (pro) of against (con) the action acting upon the theme. In example 1a, the claim would be pro as it is for the federal mandate on recycling. This is contrary to the claim in Argument 1b which is con since it is against the federal mandate on recycling. To briefly summarize, the claim of an argument includes the topic (theme) that the argument is about, as well as a verb phrase (predicate) that indicates the type of argument being made as well as the side that the author is taking. Claims are a very fundamental part of an argument, and as I will discuss in the next section, they may be vital to cue readers into the fact that they are reading an argument (Britt & Larson, 2003).

The next part of an argument to discuss is the reason. The function of a reason is to support the claim, ultimately making it more likely that the audience will accept the claim. There are many different types of reasons which can be supportive. Some reasons will appeal to values or morality. Some will simply provide data. Others may provide facts or information to increase the perception that the claim is true. The type of reason provided depends on the nature of the claim, available support for the claim, and the author’s perception of what may be most persuasive to the audience. The example arguments above use reasons which provide further

information to support the given claim. Argument 1a states that recycling should be federally mandated ‘because less material will go into our landfills.” Argument 1b states that recycling should not be federally mandated because “different types of materials can be recycled.” These reasons should support their claim and they should do so through the unstated warrant.

The warrant in Argument 1a can be stated as *‘Actions that reduce the amount of material that goes into our landfills should be federally mandated.’* Given the common background knowledge that landfills tend to be overfilling, this warrant seems minimally acceptable. So this argument would be seen as being “warranted.” The warrant in Argument 1b is ‘*actions that can be done with different types of materials should not be federally mandated.’* This warrant is more nonsensical than the one in Argument 1a. Federally mandating things that can be done with different materials seems a bit nonsensical. The warrant of this argument is unacceptable therefore this argument would be called “unwarranted.” The challenge for evaluating these warrants and teaching others this type of evaluation is that there is no single linguistic or logical test that can be applied to make this determination.

Making these kinds of evaluative judgments when reading arguments is important, but it can also be difficult for many people (Britt, Kopp, Durik, Blaum, & Hastings, 2016; Larson, Britt, & Kurby, 2009). In order to fully understand how people come to these judgments and why it is a difficult task for many, we must first discuss what we know about how people read and process arguments. The next section will discuss informal argument processing.

Argument Processing

While there is no complete model of argument processing, there are some processes that are known to be important. In the conclusion of their study, Britt and Larson (2003) lay out what they believe to be these important processes. First, in order to know that someone is reading an

argument, they must identify that a claim is being made. Once a reader recognizes that a claim is being made, their argument schema will be activated, along with prior knowledge and attitudes about the argument topic (Hample, 1977; Voss, Fincher-Kiefer, Wiley, & Silfies, 1993). This activation occurs in anticipation of an upcoming reason, which should be provided in order to add support to the claim being made. Britt and Larson (2003) found support for some of these processes. Readers read arguments faster when presented in a claim-reason order compared to a reason-claim order. This implies that when processing arguments, readers use the claim as a cue that they are reading an argument and thus activating their argument schema. Additionally, there are features of a statement that people can use to more easily recognize it as a claim (Britt & Larson, 2003). These features include qualifiers and modals. A *qualifier* is a word or phrase which changes the scope of a claim (i.e. some, all, most, etc.). For example, there is a significant difference between saying that all drugs are harmful and saying that some drugs are harmful. A modal is a type of verb which manages the scope of a verb which follows (i.e. should ban, could ban, must ban, etc.). Overall, claims and their features seem to play a special role in argument processing (Britt & Larson, 2003).

Another process that has received some support is the activation of prior knowledge and more importantly associated attitudes while reading arguments. The activation of related information during discourse processing is a well-established and researched occurrence (Albrecht & O'Brien, 1993; Kintsch, 1988, 1998; Long & Lea, 2005; Myers & O'Brien, 1998). Given that arguments are a form of discourse, we can assume that the same spreading activation of related information occurs during argument processing as well. Voss et al. (1993) conducted a series of studies in order to show that in addition to related information, our attitudes are activated while reading arguments. They based their research on Hample's (1977) model of

argumentation which modified Toulmin's structural model by adding a processing component, the activation of the reader's values/attitudes. In their first experiment, Voss et al. (1993) first asked participants how strongly they agreed with a number of statements. Two weeks later, participants were shown the statements and either made agree-disagree judgments again or were simply asked if the statement they saw was meaningful or non-meaningful, i.e. did the statement make grammatical sense. The results showed that for statements for which participants had an extreme attitude, either strongly agreed or strongly disagreed with the statement, response times were just as fast as participants who were only making meaningful judgments and these times were faster than those who held more ambivalent attitudes towards the statements. This indicates that the activation of attitudes related to a claim is just as fast as processing the text of the claim itself. This supports the idea that attitudes become activated by claims and are accessible after reading. In another study Voss et al. (1993) tested whether reading a claim would activate potential reasons. Participants were shown a claim and, after a short delay, shown a reason and then asked to decide whether or not the reason supported the claim by pressing a key. When participants agreed with a claim and the reason provided was supportive of that claim, they were much faster to respond then if they agreed with the claim but the reason provided was meant to refute the claim. This result shows that if a reader agrees with a claim they are reading, they will activate potential reasons that support the claim as they are processing the argument.

Across the studies, Voss et al. (1993) found support for their model. Both attitudes towards claims and reasons were activated upon reading a claim. Knowing that attitudes are available to readers while processing arguments begs the question of how those attitudes impact argument evaluation. The following sections will discuss argument evaluation in more detail, as well as the role that attitudes play during processing.

Argument Evaluation

People may process an argument for a variety of reasons. At times, people might be trying to learn more about a controversial topic while at other times they may skim an argument that they just happened upon while scrolling through social media. Either way, it is likely that, along with comprehending the meaning of the argument, they will also be evaluating the argument at some level. It is difficult to read an argument and not have some sort of evaluative response to it. The remainder of this section will discuss a framework of argument evaluation and research that has investigated peoples' ability to evaluate arguments.

Blair and Johnson (1987) provide a framework for argument evaluation. As Toulmin (1958) broke down arguments into three main components, Blair and Johnson (1987) break argument evaluation into three components as well. These components focus on reasons in arguments and the support they provide for their claims. The components named by Blair and Johnson are acceptability, sufficiency, and relevance.

Acceptability refers to the truth value of the reasons given in support of a claim. A reason that is acceptable is something that a 'knowledgeable', 'reflective', 'open', and 'dialectically astute' audience would agree is true. In a practical sense, it is very difficult to determine whether an audience would meet all of the criteria set by Blair and Johnson. And realistically, an audience will be composed of many people with varying levels of knowledge, openness, etc. However, it does raise an interesting question as to how these features of an audience may impact how they evaluate arguments. While this is an interesting area, it is not within the scope of this dissertation to investigate. Blair and Johnson (1987) also point out that a reason itself may not be accepted by the audience. A reason which is unacceptable may require additional support, thus acting more like an additional claim. To try to prevent this complication, I used reasons

which are statements of fact. They are minimally acceptable to a representative audience in order to eliminate differences which may come from the evaluation of truth of the claim or reason.

Sufficiency refers to the extent to which the reasons given provide enough support for the claim. Support for a claim can be sufficient if many different pieces of support are given (i.e. many reasons given to support the same claim) or if one piece of extremely strong support is given (i.e. one very strong reason is given to support a claim). For the purpose of this dissertation, I only provided one reason in support of each claim and the reason given was never extremely strong. I did not vary the number of reasons given in support of a claim. However, I did manipulate the relevance of the support given.

Relevance refers to the extent to which a reason provided is related to, or relevant to, the claim it is meant to support. Relevance is most closely tied to the warrant of an argument, the connection which links a reason and a claim. A reason that is relevant to the claim should also then create an adequate warrant between the claim and reason. By using some reasons which are not relevant to given claims, I manipulated this aspect of the arguments used in this dissertation to address Research Question 1. This manipulation is based on an evaluation task that has been used in prior research (Britt et al., 2016; Britt & Larson, 2003; Larson et al., 2009) that asked participants to make a logical quality judgment about simple arguments as to whether the argument was logically or structurally flawed. All versions of this flawed judgment task included as low-quality arguments, unwarranted arguments in which the reason was not relevant to the claim. In some versions of this task, low-quality arguments also included unsupported assertions (claims without reasons). In the current dissertation, I used unwarranted arguments for the low-quality arguments. Collecting reading time data while using an evaluation task, like the flawed

judgment task, allowed me to investigate how readers process these arguments when evaluating them for the relevance of the reason provided.

To ground this, let's look at back at arguments 1a and 1b from earlier:

> **1a: Pro-Warranted.** Recycling should be federally mandated because less material will go into our landfills.
>
> **1b: Con-Unwarranted.** Recycling should not be federally mandated because different types of materials can be recycled.

If we look at Argument 1a and evaluate the acceptability of the argument, we find that the reason provided is acceptable, meaning it is true. A reasonable audience would agree that our landfills would become smaller if we recycled more. Similarly, the reason provided for Argument 1b is also acceptable. It is true that different types of materials can be recycled. Additionally, both items are equally insufficient and incomplete in that there is only one reason given in support of each and the reason provided is not extremely strong. But, the arguments do differ in relevance. Argument 1a has a reason that is relevant to the claim. A reduction in the amount of material going into landfills is related to whether or not recycling should be mandated. Argument 1b, on the other hand, does not have a relevant reason. The fact that different types of materials can be recycled has no bearing on mandating recycling. While it is semantically related, it is not relevant to the claim being made. Moving forward, I will call arguments with relevant reasons as *warranted* and arguments with irrelevant reasons as *unwarranted*. So, Argument 1a is warranted and arguments 1b is unwarranted.

Research on argument evaluation has typically focused on relevance as this type of evaluation directly maps onto the warrant in Toulmin's argument structure. The warrant is the piece of an argument that must be inferred by the reader which leads to many different

opportunities for investigation. This research has aimed to understand how to improve evaluation skill (Larson et al., 2009), identify features of those who are skilled at evaluation (Britt et al., 2016), and other features of arguments which may impact evaluations (Dandotkar, 2012).

Larson, et al. (2009) tested the effectiveness of a set of tutors on improving the skill of evaluating the logical quality of simple informal arguments. By logical quality, Larson et al. (2009) are referring to whether an argument is warranted or unwarranted. A higher quality argument would have a relevant reason and be warranted where a low-quality argument would have an irrelevant reason and be unwarranted. In the first experiment participants were given a flawed judgment task. This test was comprised of 24 simple arguments such as 1a and 1b. One third of the arguments were warranted (e.g., 1a), meaning the data provided to support the claim was relevant. Another third of the items were unwarranted (e.g., 1b), the data provided was not relevant to the claim. The final third of the items were unsupported assertions. Items in this condition were simply claims with no data presented. Participants were instructed to decide whether the arguments provided were logically or structurally flawed (unwarranted or warranted). Participants were also told to make the judgments only considering the structure of the arguments and avoid allowing their agreement or persuasiveness of the argument to affect their judgements. This study found that without training participants were only 66% accurate at evaluating the arguments for relevance (with chance performance at 50%). A 15-minute tutorial, which trained participants to focus on the features of arguments to better evaluate them helped the participants reject unsupported arguments, but did not help participants distinguish between warranted and unwarranted arguments. A second experiment made alterations to the tutor in order to present it on a computer and provide feedback. Using the same flawed judgment task as Experiment 1, the researchers found that those given immediate feedback from the computer

version of the tutor were able to significantly improve their ability to reject unwarranted arguments, thus overall improving their ability to evaluate arguments for the relevance of the reason provided. While Larson et al. (2009) were able to use a tutor to improve this kind of evaluation, there was still a room for improvement. This type of evaluation may not only be impacted by the participants' skill at evaluating arguments, but also features of the arguments themselves. Dandotkar (2012) investigated whether less skilled readers differed from high skilled readers in how semantic overlap across an argument impacted argument evaluations on a flawed judgment style task and an agreement task.

Dandotkar (2012) examined the relationship between semantic overlap and reason relevance (Blair & Johnson's relevance) on participants' quality judgments and agreement judgments. Semantic overlap is the extent to which word meaning coexists in discourse. For example, The U.S. should legalize marijuana because marijuana grows naturally in the U.S., has high semantic overlap. The terms "U.S." and "marijuana" are present in both the claim and reason of the argument. A low semantic overlap example would be, The U.S. should legalize marijuana because it grows naturally here. There are no key thematic words that are present in both the claim and reason. Dandotkar (2012) hypothesized that either semantic overlap would influence evaluations (semantic dominance) or that reason relevance would impact argument evaluations (logical dominance). Dandotkar (2012) found support for the logical dominance hypothesis as the results showed that arguments that were warranted were judged higher than those which were unwarranted. Essentially, the logical connection between the claim and reason in the given arguments (the warrant) was evaluated during both quality and agreement judgments. This was true for both high and less skilled readers. This indicates that less skilled reasoners are not mistakenly using semantic overlap to accept unwarranted arguments. In fact, it

was the skilled reasoners that used high semantic overlap to increase their judgments of warranted arguments. These results show that readers are sensitive to reason relevance but that their difficulty in discriminating warranted from unwarranted arguments is due to something other than simple word or semantic overlap.

These studies show that while many people are able to evaluate arguments for their logical quality, it is still a difficult task for most. Even those who are proficient at argument evaluation still aren't perfect. There seems to be something at play that makes it difficult for people to attain true mastery at this skill. One possibility, and the focus of this dissertation, is that prior attitudes about arguments may somehow impact these types of evaluations under some conditions.

Attitudes and Beliefs

The next section will focus on research that has tried to examine the role that prior attitudes and beliefs play in argument evaluation. Before discussing this literature too deeply, I want to clarify some terms related to attitudes and beliefs.

When discussing the role of attitudes in argument evaluation, it is important to first understand what is meant by an attitude and how it differs from a belief. On the surface, they seem like very similar concepts, in fact, even in the field of social psychology, where attitude research is common, there can be disagreement or confusion. For example, in this dissertation I used a scale of attitude certainty created by Petrocelli, Tormala, and Rucker (2007). This scale was created to be used to measure attitude certainty. Though, the first sentence of this paper states, "Psychological certainty is the cornerstone of beliefs" (Petrocelli, et al., 2007, pg. 1). While there is some confusion about the exact nature of attitudes and beliefs, Wolfe and Griffin (2016) tried to clarify the distinction.

According to Wolfe and Griffin (2016), beliefs represent someone's conclusion about the truth value of about a topic or idea whereas attitudes represent someone's affective evaluation of a topic or idea (Wolfe & Griffin, 2016). I think this distinction can be seen with the following example. Imagine reading a weather report that states the sky is green. This likely goes against your knowledge that the sky is blue. As a result, you do not believe the sky is green, so you conclude that it is not true. Now, imagine reading a blog that proposes that we should change the color of the sky to be green. Whether or not you agree with the blogger about changing the color of the sky to green would represent your attitude about the topic. This is more of a preference and is not something that can necessarily be known to be true or not.

Given the nature of the proposed study, I will be discussing attitudes instead of beliefs. That being said, I will still be using research that discusses beliefs when it is relevant to the current project.

Attitudes and Argument Evaluation

As mentioned before, an argument with a reason that does not act as support to the argument are unwarranted and should be illogical to those that hear it, but this is not always the case. This could be because readers may not evaluate the arguments they encounter solely on the relationship between the claim and reason. Readers, instead, may allow their attitudes about the topics involved in the argument to play a role in their evaluation.

Take the following arguments.

2a. The U.S. should ban bump stocks because they make mass shootings more deadly.

2b. The U.S. should ban bump stocks because they are accessories for firearms.

Both Argument 2a and 2b have the same main claim that the U.S. should ban bump stocks. They differ in the reason provide to support that claim. Argument 2a provides a

warranted argument, meaning there is a logical connection between the claim and reason. If bump stocks do make mass shootings more deadly, it makes sense to ban them in order to prevent that. On the other hand, Argument 2b is unwarranted, there is no direct logical connection between the claim and reason. Just because something is an accessory to a firearm doesn't mean it should get banned. Being an accessory to a firearm doesn't support the claim that bump stocks should be banned. But as mentioned, people do not always evaluate arguments in this logical way, sometimes their attitudes towards the topic can impact how the arguments are evaluated.

For example, if I am a member of the NRA, a gun enthusiast, or believe in small government, I may say that both arguments are bad, simply because I disagree with the claim and I may spend more time processing these arguments, scrutinizing them and looking for flaws other than simply the relevance of the reason since I disagree (Edwards & Smith, 1996; Lord, Ross, & Lepper, 1979).

Social Psychology Perspective

Edwards and Smith (1996) proposed the Disconfirmation model of argument evaluation that incorporates the readers' prior beliefs as an important part of the evaluation process. While they discussed their research in the context of beliefs, 6 of the 7 "belief statements" were "should/should not" and the other was "is/is not appropriate". Thus, the work would now be considered to be on attitudes, not beliefs. This model, shown in Figure 1, posits that when someone encounters information that they do not agree with, they will spend time scrutinizing the argument by actively searching their memory for information to disconfirm the argument.

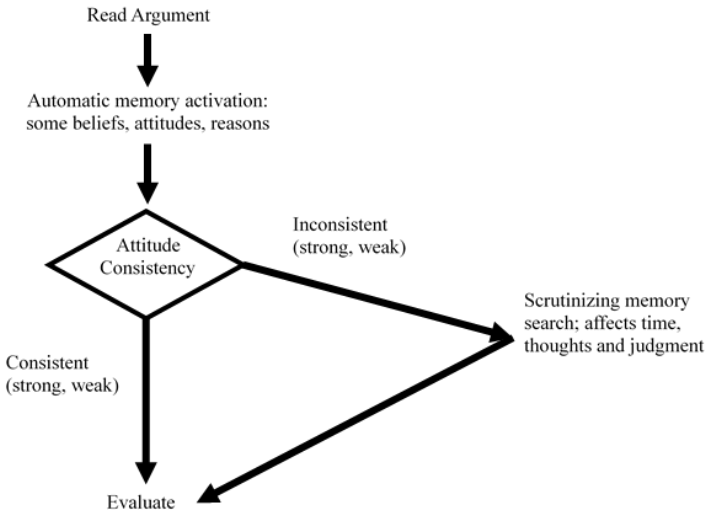

Figure 1. Disconfirmation Model proposed by Edwards and Smith (1996).

Edwards and Smith (1996) found support for this model across two experiments in which they had participants read short reason-claim arguments and measured strength evaluations, reading times, and typed think-aloud responses. They found that arguments that were inconsistent with their prior attitudes for the claim were evaluated as less strong than those that were consistent with their prior attitudes. This result directly supports the idea that attitudes can impact this type of argument evaluation. Additionally, participants spent more time reading arguments that they disagreed with compared to those which they did agree with. This suggests that there is some additional processing going on when people are reading arguments that they do not agree with. Finally, when presented with arguments incompatible with their attitudes, participants produced more refutational statements about that topic compared to those that were consistent with their prior attitudes. This pattern of findings suggests that the processing that occurs is effortful and is directed to scrutinize the attitudes inconsistent arguments. Taken together, these results do provide fairly strong support for their model.

An important thing to note is that Edwards and Smith (1996) asked participants to evaluate the strength of the arguments they read. While this type of evaluation seems to differ

from evaluations of relevancy (like the flawed judgment task) it is actually very similar. The strength evaluation used by Edwards and Smith (1996) asked participants to judge the extent to which the premises provided (reasons) lead to the given conclusion (claim). This task also explicitly told participants that their only job was to evaluate this logical relationship and that they should not evaluate whether they think the conclusion is true or false or whether they agree or disagree with the argument. This type of task is very similar to the strength task used by Shaw (1996) which also asked participants to evaluate how well the premises support the conclusion not whether they agreed with the premises and conclusion. This task is also very similar to the flawed judgment task used in previous studies (Britt et al., 2016; Britt & Larson, 2003; Dandotkar, 2012; Larson et al., 2009) that asked participants to decide whether the arguments were logically or structurally flawed and were told to not consider their agreement with the argument or its persuasiveness.

Building upon the model proposed by Edwards and Smith (1996), Clark and Wegener (2013) propose the Discrepancy Motives Model shown in Figure 2. The Discrepancy Motives Model was proposed to explain the set of conditions in which people may undergo effortful processing of arguments. While Edwards and Smith's Disconfirmation model of argument evaluation focuses on the scrutiny and close examination of attitude inconsistent arguments, the Discrepancy Motives Model proposes that there are circumstances in which people engage in deeper processing of arguments that may be attitude consistent as well. The model posits that when people encounter attitude inconsistent information that they hold strong attitudes about, they are likely to attempt to defend their position and will scrutinize the information. This is in line with the predictions of the Disconfirmation Model (Edwards & Smith, 1996).

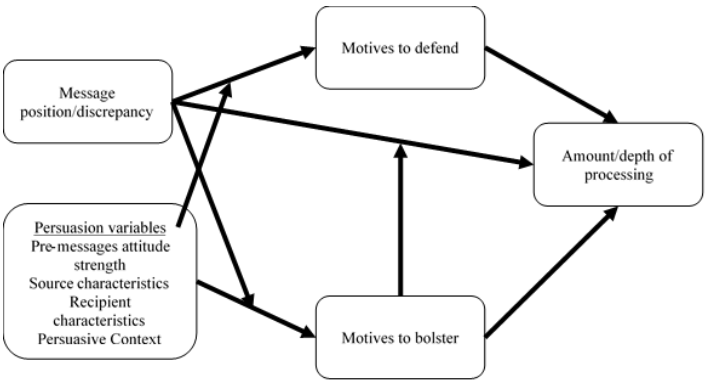

Figure 2. The Discrepancy Motives Model as proposed by Clark and Wegener (2013).

In addition, the Discrepancy Motives Model expects that when people hold weak attitudes about a topic, they may attempt to strengthen, or bolster their beliefs if they hear arguments that are consistent with their attitudes about that topic. One aspect that may lead to bolstering is attitude accessibility (Clark et al., 2008). Attitude accessibility is essentially how quickly and strongly attitudes about a topics become activated. In their study, Clark et al. (2008), examined whether this factor could affect post-reading attitudes after reading an extended argument with four reasons. They found that when arguments were attitude inconsistent, increasing attitude accessibility was related to an increase in processing. This is consistent with the scrutiny effect which they refer to as defending one's position. They also found that when arguments were attitude consistent, increasing attitude accessibility was related to a decrease in processing.

This second finding is evidence for bolstering. Suggesting that when encountering information which supports an attitude that is weakly held, people will use it to strengthen their attitudes. This model works as an extension of the model proposed by Edwards and Smith (1996), and was created to help understand how attitudes interact with argument processing.

The Discrepancy Motives Model of 2013 provides a larger framework for understanding when motivated defending or motivated bolstering will occur in persuasive contexts. In Figure 3,

I present a simplified model to remove the other set of persuasive factors (e.g., the source of the message) and to map it onto the process model format of Edwards and Smith (1996). This shows the added bolstering process.

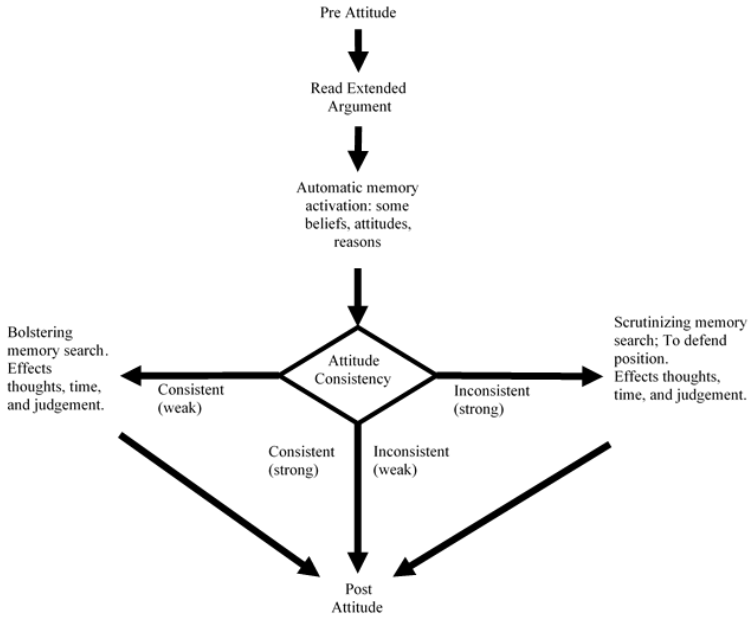

Figure 3. Modified processes model based on Clark and Wegener (2013) two motivated processes for argument evaluation.

Much of the work on attitudes related to argumentation has been conducted in social psychological research but there is some research related to these ideas in cognitive psychology as well.

Cognitive Psychology Perspective

While Dandotkar (2012) did not directly address the effect of attitudes on flawed judgement task performance in his study, he did conduct a post hoc analysis. He found that participants' attitude about a topic was also related to their quality judgments. Participants were 4.36 times more likely to accept attitude consistent arguments compared to attitude inconsistent arguments.

Wolfe and Kurby (2017) investigated the extent to which belief in the claim of an argument influences argument evaluation and whether differences in reasoning ability and the importance of the topic to the reader had an impact on this effect. In the first experiment, Wolfe and Kurby (2017) collected participants' beliefs about the topic of spanking as a punishment for children and whether watching television violence causes people to act violently in real life. Participants were then asked to evaluate 20 arguments about spanking (10 pro, 10 con), 20 arguments about TV violence (10 pro, 10 con) and 20 filler items for their logical quality. This task was similar to the flawed judgement task used by Larson et al. (2009). The results showed that beliefs about the topic did not impact overall accuracy in evaluating arguments. Whether or not a participant believed the claim of the argument they were evaluating did not influence their evaluation of the argument's quality. However, participants did show a bias to accept arguments that they agreed with and rejected arguments that they did not. In their third experiment, Wolfe and Kurby (2017) found that reasoning ability did not reduce the response bias, but good reasoners were overall more accurate at evaluating arguments than poorer reasoners. This is also consistent with what Dandotkar (2012) found.

Issues with Prior Research

While these studies have helped us better understand how attitudes impact our evaluation of arguments, there are still many unknowns to investigate.

Edwards and Smith's (1996) findings were impactful at the time but some of their methods can be improved upon or extended. Firstly, in their study they used a preface-conclusion model of argumentation instead of a claim-reason model. In doing so, they presented participants with the prefaces first, which are akin to reasons. Research shows that claims hold an important place in argument processing (Britt & Larson, 2003). As such, I first attempted to replicate their

key results using the more canonical claim-reason order. Additionally, Edwards and Smith (1996) used a task which potentially conflated argument strength with argument quality, the former of which lends more to subjective interpretations where attitudes may play a larger role. I tried to remedy this by giving participants different types of tasks while reading the arguments. Finally, Edwards and Smith did not use unwarranted arguments. All of the arguments presented were at least minimally supportive to the claim. Using completely unwarranted arguments in the current study allowed me to better test how attitudes impact processing and evaluation, especially since the primary task of interest is an evaluation task. Neither Edwards and Smith (1996) nor Clark and Wegener (2013) directly examined whether strength of attitudes matters to the effect. Voss et al. (1993) found that strong beliefs lead to faster activation of related beliefs than moderate beliefs. Given this result, it is important to test how strength of attitude may impact argument evaluation and processing.

I use the terms warranted and unwarranted as in prior work (Larson et al., 2009). The warranted arguments were "minimally warranted" in that the connection between the reason and the claim forms a generalized associated conditional statement (Gilbert, 1999) that would be acceptable to a reasonable audience without additional inferences. For example, in Argument 1a, the warrant is that *Actions that reduce the amount of material that goes into our landfills should be federally mandated.* In general, things that reduce our landfills would be a good thing. In contrast, unwarranted arguments are not directly acceptable without inferences that create a different warrant. For example, in Argument 1b, the warrant is that *Actions that can be done with different types of materials should not be federally mandated.* In this case, being able to complete an action with different types of materials is not a minimally acceptable situation for requiring the behavior. These unwarranted arguments are interesting to examine because the reader could

do additional inferencing to make the argument warranted and attitude consistency is one motivator for that inferential work.

Cognitive psychological research on argumentation and attitudes has been limited as well. Some of the research on argumentation does not consider attitudes at all (Larson et al., 2009). Dandotkar (2012) did find that attitudes may impact argument evaluation, but it was found as a post-hoc analysis and the study wasn't designed to ask or answer that question. Wolfe and Kurby (2016) most directly addressed the role of beliefs, not attitudes. However, their studies used 40 arguments about only two topics. This may lead to large contrast effects across items. The arguments used were also causal in nature, instead of policy based, which may change processing. They did a classic flawed judgment task with a binary judgement on argument acceptability and did not find an effect of belief consistency on judgments. But it is possible that a binary judgment may not be sensitive to finding these kinds of effects on judgements. That doesn't mean there will not be an effect of attitudes on argument evaluation since policy claims should have a higher bar set then claims of behavior changes (e.g., "should") and Likert scales may be more sensitive and able to capture the effect.

Another important note, much of the research on argument evaluation from a cognitive perspective intentionally chose topics that were not particularly controversial (i.e. recycling, speed limits) (Britt & Larson, 2003, Larson et. al, 2009). This allows the research to focus on processing but ignores more realistic situations in which one might encounter an argument. On the other hand, research from a social psychology perspective often uses more controversial topics (i.e. death penalty, abortion etc.) (Edwards & Smith, 1996). This, along with measurements of argument strength tend to embrace the impact of attitudes on evaluation but doesn't account for processing of less controversial arguments.

Overview of Experiments

This dissertation extended previous research to address the boundary conditions for the attitude consistency effects. Prior research has shown that adults can detect unwarranted arguments, but this skill of evaluating logical quality is not consistent nor at the level (65% accurate) that we would describe as mastery (80% accurate) (Larson et al., 2009). It would be important to examine whether attitudes lead one to accept clearly unwarranted arguments and whether that helps explain why participants in these studies do not evaluate arguments for their logical quality at mastery level, even after training. One possibility is that people give the arguer the benefit of the doubt for attitude-consistent unwarranted arguments and as a result make the necessary inferences to make the argument seem warranted. It is also important to use unwarranted arguments to show that participants are actually processing the arguments while evaluating. In order to successfully evaluate unwarranted arguments as less logically sound, participants must be processing those arguments with some effort. Although Wolfe and Kurby (2017) used unwarranted "bad" arguments, they used only two topics and the arguments used had causal claims (which may be harder to evaluate than policy claims). This dissertation used a greater number of topics to increase generalizability and all arguments had policy claims, to stay consistent with most prior research.

The goal of this dissertation was to investigate boundary conditions of argument evaluation that have not yet been explored. The basic design assessed attitudes toward a set of claims and then in a later session had participants evaluate argument quality. The main within participant factors were the quality of the argument (warranted vs. unwarranted) and whether it is consistent with the participant's attitude (consistent vs. inconsistent). In Experiment 1, there was also a task manipulation to compare instructions for evaluation, either quality (logical

evaluation) or strength (Edwards & Smith, 1996). In Experiment 2, the task manipulation was evaluation for quality vs a non-evaluative control task. I will present the research questions first and then present the hypotheses and predictions after the research questions.

Research Question 1

Is there a boundary condition on attitude consistency effects related to the materials? Specifically, is argument acceptability a boundary condition to the attitude consistency effect on argument evaluation? To answer this question I will manipulate the quality of arguments and compare the processing and evaluation of these warranted and unwarranted arguments. This manipulation allows me to test whether the attitude consistency effect is constrained by the Logical Dominance Hypothesis (below). Although this hasn't been tested with unwarranted arguments, there are two ways in which comparing warranted to unwarranted arguments could affect the amount of scrutiny. There could be more scrutiny on arguments that one disagrees with that are minimally acceptable (warranted). This would show an interaction on evaluations and reason reading time. Or, more scrutiny on those arguments that one disagrees with regardless of logical quality (attitude consistency effect on warranted and unwarranted arguments) (see Disconfirmation Hypothesis below). This would be consistent with Wolfe and Kurby (2017) who found a response bias for being more likely to accept unwarranted arguments that one agrees with.

Research Question 2

Is there a boundary condition on the attitude consistency effect related to the reader? Specifically is attitude strength a boundary condition to argument evaluations due to the attitude consistency effect?

Edwards and Smith (1996) tested attitude strength as a categorical variable with two extremes (strongly disagree and strongly agree. However, they found that it was likely the effect would have help across all attitude strength levels. On the other hand, Voss et al. (1993) found that extreme attitudes are processed differently than more neutral attitudes. Similarly, Clark et al. (2008) used attitude accessibility as a proxy for attitude strength and found that processing of attitude inconsistent arguments is increased as attitude accessibility increased. This pattern was reversed for attitude consistent arguments. To answer this question I will ensure participants vary in their attitude strength towards the topics of argument (strongly disagree, weakly disagree, weakly agree, strongly agree). This manipulation allows me to test whether the attitude consistency effect is due to scrutiny for all attitude strengths (Disconfirmation Hypothesis) or if for some strength levels there is bolstering as well (see The Bolstering and Defending Hypothesis below).

Research Question 3

Is there a boundary condition on the attitude consistency related to the task? Specifically is type of task a boundary condition to argument evaluations due to the attitude consistency effect? To answer this question in Experiment 2 I will include a non-evaluative task. It is possible that the attitude consistency effect will only occur during more evaluative tasks which elicit effortful processing (i.e. scrutiny). However, Voss et al. (1993) found that extreme attitudes activated related information as quickly as comprehension. Although extreme attitudes are activated quickly, they may only influence processing when a task requires evaluation. This manipulation of task allows me to test whether the attitude consistency effect is constrained by the Evaluative Mindset Hypothesis (below).

Hypotheses and Predictions

To examine these research questions, there are four hypotheses that can be tested as possible mechanisms.

The *Logical Dominance Hypothesis* states that when presented with arguments, participants should rely solely on the logical connection between the claim and reason to make evaluations (Dandotkar, 2012; Larson et al., 2009) (Experiment 1 and 2). This hypothesis predicts that participants will rate warranted arguments as higher quality than unwarranted arguments (Experiment 1 and 2).

The *Disconfirmation Hypothesis* states that when presented with attitude inconsistent arguments readers will engage in scrutiny in order to find flaws with the argument to defend their attitude (Edwards & Smith, 1996) (Experiment 1 and 2). This hypothesis predicts longer reading times for attitude inconsistent arguments compared to attitude consistent arguments. Additionally readers will rate attitude consistent arguments as higher quality compared to attitude inconsistent arguments.

The *Bolstering and Defending Hypothesis* states that when presented with attitude inconsistent arguments about topics which readers strongly disagree with, readers will engage in scrutiny (defending) in order to find flaws with the argument to defend their attitude. Additionally, when readers are presented with attitude consistent arguments about topics with which readers weakly agree, they will engage in bolstering in order to strengthen their already existing, but weak attitudes (Clark & Wegener, 2013) (Experiment 2).This hypothesis predicts that readers will spend more time processing arguments that they strongly disagree with, in order to defend their position, and also for arguments that they weakl agrees with, in order to bolster their position, compared to other arguments.

The *Evaluative Mindset Hypothesis* states that when placed in an evaluative mindset, argument processing will be influenced by activated information, including prior knowledge and attitudes, in order to come to an evaluative conclusion (Britt & Larson, 2003) (Experiment 2). This hypothesis predicts that readers given an evaluative task while reading arguments will spend more time processing arguments than those given a non-evaluative control task.

CHAPTER 2: EXPERIMENT ONE

In this experiment, I examined participants' evaluation of arguments with policy claims. In order to understand how attitudes impact argument evaluations I first measured participants' attitudes about 15 claims. Later, participants evaluated four arguments created from each claim: two pro and two con versions supported by a warranted and an unwarranted reason. Based on the disconfirmation model, attitude inconsistent arguments should be evaluated as less strong than attitude consistent arguments. Based on the logical dominance hypothesis, warranted arguments should be evaluated as stronger than unwarranted arguments.

Methods

Participants

Forty-three undergraduate students (59% female) from two low level general education classes in the Psychology department at a Midwestern university were selected that had at least 2 claims that they rated as strongly disagree (1-3), disagree (4-5), agree (6-7), and strongly agree (8-10) in an initial session at least 2 weeks prior to the experiment. The instructions for this task can be found in Appendix A.

Design and Materials

The design was a 2 Evaluation task (quality vs. strength) between participants x 2 Quality (warranted vs. unwarranted) within participants x 2 Attitude Consistency (consistent vs. inconsistent) within participants design. For each of the 15 claims, four separate arguments were created (resulting in 60 claim-reason arguments). There were 2 pro arguments and 2 con

arguments. For each of these claims, a reason was selected to create a warranted argument (reason is true and provides support for claim) vs. unwarranted (reason is true but does not provide support) argument. A full list of arguments can be found in Appendix B. Half of the participants were randomly assigned to rate the quality ("For each issue, you will be asked to evaluate the logical quality of each argument. By 'logical quality' we mean whether the reason could provide support for the claim to a general audience. Thus, your job is to judge extent to which the reason is structurally relevant to the claim.").The other half rated the strength ("For each issue, you will be asked to evaluate the strength of each argument. By 'strength' we mean how convincing the argument would be to a general audience. Thus, your job is to judge how strongly the reason makes the claim more believable -- NOT whether or not you agree with the argument.). There was no difference of evaluation task on any of the measures, so the rating task will not be discussed further.

Procedure

At least 2 weeks after rating their attitudes on the 15 claims, participants were invited to participate in the evaluation portion of the study using Qualtrics. The 60 arguments were presented to participants one at a time using block randomization. Each block contain an argument about each topic and an equal number of pro, con, warranted, and unwarranted arguments. Participants were shown an argument and asked to rate either the logical quality or strength of the argument (depending on condition) on a 10 point Likert scale. After making a judgment on one argument participants pressed a "next" button and the next argument was shown on the screen.

Data Transformation

After completion of the study, data was organized using the participants' responses to the attitude measure in the pretest survey. This allowed responses to pro and con arguments to be categorized as responses to attitude consistent or attitude inconsistent arguments. For example, if a participant were to rate their attitude about banning late term abortions as "strongly agree", then their responses to the pro "banning late term abortions" arguments was categorized as responses to attitude consistent arguments. Both the argument and the participant and argument were for banning late term abortions. On the other hand, their responses to con arguments about banning late term abortions were categorized as responses to attitude inconsistent arguments. In this case the participant was strongly for banning late term abortions but the argument was against banning late term abortions. If a different participant "weakly disagrees" with banning late term abortion, then this pattern was reversed. Since this person is against banning late term abortion, any arguments that are for banning late term abortion would be attitude inconsistent whereas arguments against banning late term abortion would be attitude consistent. Table 1 shows an example of expected evaluations based on this transformation. Participants were selected who had at least two responses in each of the four agreement categories.

Table 1

Four sample argument strength ratings and the corresponding attitude item by attitude consistency and idealized rating based on attitude for minimally acceptable arguments only.

	Argument Strength Rating			
Experiment 1	**Strongly agree**	**Weakly agree**	**Weakly disagree**	**Strongly disagree**
Attitude item The U.S. should ban late term abortions	**8 to 10**	**6 to 7**	**4 to 5**	**1 to 3**
Argument: Consistent with side of item The U.S. should ban late term abortions because they can save the lives of some women.	10	7	4	1
Argument: Inconsistent with side of item The U.S. should *not* ban late term abortions because they can save the lives of some women.	1	4	7	10
Experiment 2	**Strongly agree**	**Weakly agree**	**Weakly disagree**	**Strongly disagree**
Attitude item The U.S. should ban late term abortions	**4**	**3**	**2**	**1**
Argument: Consistent with side of item The U.S. should ban late term abortions because they can save the lives of some women.	8	6	3	1
Argument: Inconsistent with side of item The U.S. should *not* ban late term abortions because they can save the lives of some women.	1	3	6	8

Results

A 2 Quality (warranted vs. unwarranted) x 2 Attitude consistency (consistent vs. inconsistent) X 4 Attitude Strength (strongly disagree, disagree, agree, and strongly agree) within participants ANOVA was conducted on ratings (10 point scale). The main effects for all three factors were significant. Importantly, participants rated attitude consistent arguments (M=5.55, SD=0.76) higher than attitude inconsistent arguments (M=4.76, SD=0.87) ($F(1,43) = 41.83$, $p <$

.001) and they rated warranted arguments (M=6.43, SD=0.98) higher than unwarranted arguments (M=3.88, SD=1.03) ($F(1,43) = 140.45$, $p < .001$). The main effect of attitude strength was significant as well, $F(3, 129) = 3.06$, $p = .031$. While there was an attitude consistency effect, it was smaller than the quality effect showing that attitude consistency does not make it impossible to evaluate quality. There were also two significant interactions: Agreement X Attitude Consistency ($F(3,129) = 31.00$, $p < .001$) and Quality X Attitude Consistency ($F(1,43) = 14.46$, $p < .001$).

As shown in Figure 4, the attitude consistency effect is only significant for warranted arguments. Thus, participants rate warranted arguments higher when they are consistent with their attitudes but when the arguments are unwarranted, there is no effect of their attitudes on their ratings.

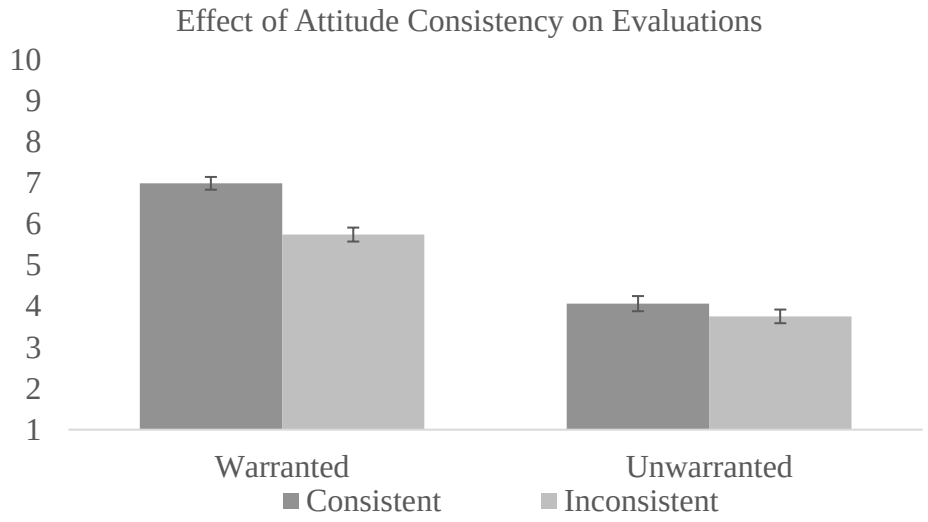

Figure 4. Effect of Attitude Consistency on Evaluations. Error bars (SE).

As shown in Figure 5, the attitude consistency effect is only significant for claims that they agreed with or strongly agreed with. In fact, there is a trend (p=.06) for the opposite

direction for the claims that they disagree with. Attitude strength is based on participants' initial attitude agreement judgements.

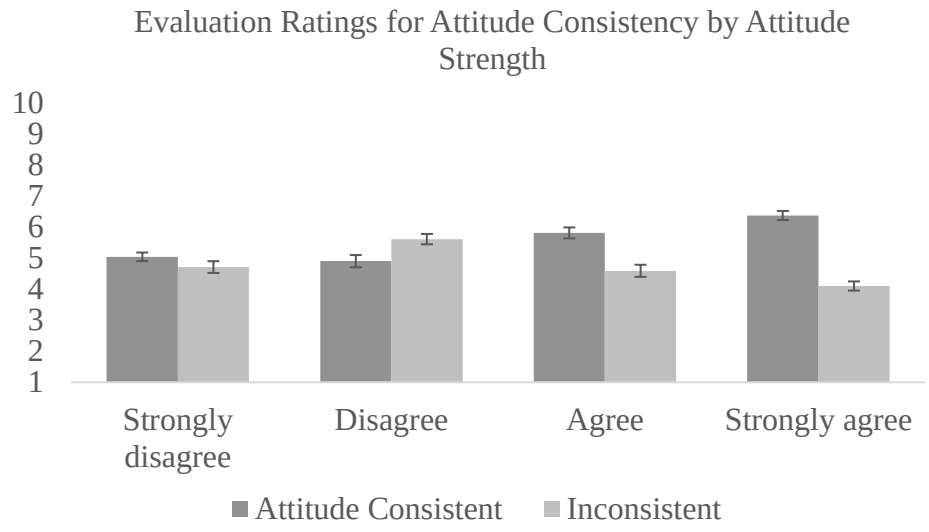

Figure 5. Evaluation Ratings for Attitude Consistent/Inconsistent Arguments for Each Agreement Category.

Discussion

Overall this study shows that a person's attitudes about everyday topics impact how they evaluate arguments they may read about those topics. I did replicate the attitude consistency effect; however, I also found a limit based on the Logical dominance hypothesis. Participants' attitudes did not override one's skill at judging logical arguments. This logical dominance was also supported in terms of the interaction of attitude consistency and quality. The attitude consistency effect was mostly found for arguments that one agrees with and leads to higher ratings of warranted arguments for those claims that you agree with rather than increasing the rejection of unwarranted argument for claims you disagree with. Though, this effect could be due to the phrasing of the items in the pretest survey. In this study, 12 of the 15 items were framed as

pro claims (i.e. we should ban) while only 3 were framed as con claims (i.e. we should not ban). This limitation is something that will be addressed in Experiment two.

The results of Experiment 1 replicate some of Edwards and Smith's (1996) findings. An attitude consistency effect showed that attitude-consistent arguments were rated higher than attitude-inconsistent arguments. Interestingly, I found three constraints on this effect. Firstly, the attitude consistency effect did not override readers' ability to evaluate argument quality. Secondly, this effect was only found for warranted arguments. Meaning, participants did not simply reject unwarranted arguments that they didn't agree with more than unwarranted arguments they did agree with. Rather, participants rated warranted arguments that they did agree with higher than warranted arguments they did not agree with. Finally, the impact of attitudes on evaluation was limited to arguments with claims that participants agreed with, rather than arguments with claims that they disagreed with. These findings help give some answers to my research questions. Both argument quality and attitude strength impacted evaluations. This suggests that they may act as boundary conditions to the attitude consistency effect. Also, it seems that unwarranted arguments are recognizable and dismissed without an impact of attitudes on evaluation whereas logically relevant arguments are prone to being impacted by attitudes during evaluation. Whether or not this pattern holds true during processing is a question for Experiment 2.

CHAPTER THREE: EXPERIMENT TWO

The results of Experiment one found that a person's attitude about a topic can impact their evaluations of arguments about that topic. To better understand why the effect of attitudes on argument evaluations, Experiment two measured both participant evaluations and two different measures of processing: reading time and thought listing responses. By measuring both processing time and thought listing responses I hope to be able to understand how differences in processing may lead to differences in final evaluations of arguments. However, since thought listing responses were only be done on a subset of the arguments, it was not be a primary dependent variable. I also be investigated the nature of the role that attitudes play during evaluation. While attitude strength ratings give insight into the intensity which one holds an attitude about something, there may be other features of attitudes that can impact evaluation and processing as well. One such feature is *attitude certainty*. Attitude certainty is the extent to which someone is sure that they know what their attitude is and how confident they are that their attitude about a topic is the correct attitude to hold (Petrocelli et al., 2007). Although it was my intention to use attitude certainty as covariates, after data loss, there were not enough participants, so attitude certainty was not included in the analyses. Experiment 2 also investigated further the relationship between attitude strength, argument evaluation, and processing.

Additionally, Experiment 2 included a condition where participants do not evaluate arguments. This condition allowed me to compare argument processing without an underlying

evaluation task to processing with an evaluation task. It could be that being engaged in an evaluation task may foster the processing differences found in prior research.

Methods

Participants

Data was collected from 141 Mechanical Turk workers for part 1 of the study (average age: 38.59 years; 66% Male, 34% Female; 72% Caucasian, 12% Asian, 8% Hispanic, 7% African American). Data was collected from 69 participants for part 2 of the experiment. Only data from 42 of those participants were used for analyses (average age: 40.24 years; 56% male, 42% female; 71% Caucasian, 14% Asian, 12% African American, 2% Hispanic). Originally data was going to be collection using the same population as Experiment 1, but due to the global Covid-19 pandemic, data collection had to be changed to online using a different population.

Data Loss

Of the initial 141 participants who completed part 1 of the experiment only 69 returned to complete part 2. Of those 69, 6 were removed for non-compliance in part 1 (reading and responding to instructions and 48 items in less than 2 minutes). Of the remaining 63, 6 were removed for non-compliance in part 2 (responding to more than 20% of items in less than half a second). Finally, of the remaining 57, 15 did not meet criteria to be included in the experiment. These participants did not have varied enough attitudes across topics to use their data to test the effect of attitude strength. Data from the remaining 42 participants was analyzed in the experiment.

Design and Materials

The experiment was conducted as a 2 Quality (warranted vs. unwarranted) within-participants x 2 Consistency (consistent vs. inconsistent) within-participants x 2 Task (evaluate vs. control) between-participants, mixed methods design.

Pretest Survey

Ten to fourteen days prior to completing the main study, attitudes for each claim (Edwards & Smith, 1996), certainty of those attitudes (Petrocelli et al., 2007) and prior knowledge about the topics (Edwards & Smith, 1996) were measured. This survey was constructed in Qualtrics and administered through Mechanical Turk. Participants were compensated $0.50 for completing this part of the study.

Attitudes

Attitudes were measured similarly to Edwards and Smith (1996). This measure asked one question about the participants' attitudes per topic by asking participants to rate the extent to which they agree with a statement. For example, for the topic of marijuana legalization, the question may say, "The U.S. should legalize recreational use of marijuana." Participants then select a number between 1 (strongly disagree) to 4 (strongly agree). I chose a 4 point scale to replicate the measure used by Edwards and Smith (1996) as well as to force participants to choose a side by using a scale without a middle point. The instructions of this measure can be found in Appendix D.

Attitude Certainty

Certainty about attitudes was measured using a modified scale created by Petrocelli et al. (2007). This measure breaks attitude certainty into two parts, attitude clarity and attitude correctness. According to Petrocelli et al. (2007), attitude clarity is "the subjective sense that one

knows what one's attitude is" and attitude correctness is "the subjective sense that one's attitude is correct or valid" (Petrocelli et al. 2007). According to the authors, these constructs together give a measure of attitude certainty. Through their validation study, the authors ended with a 7 item survey, four items to measure attitude clarity and three items to measure attitude correctness. The full survey can be seen in Appendix E. For the sake of limiting participant fatigue, I chose one item for attitude clarity and one item for attitude correctness.

Attitude clarity: "To what extent is your true attitude toward TOPIC clear in your mind?"

Attitude correctness: "How certain are you that your attitude toward TOPIC is the correct attitude to have?"

Participants answered each question about all 16 topics. Due to lack of power resulting from a small sample size, this measure was not used in any analysis.

Prior Knowledge

Prior knowledge was measured using the same measure as Edwards and Smith (1996). This measurement is on a 4 point Likert scale ranging from 1 (I have no knowledge about this topic) to 4 (I have a great deal of knowledge about this topic). The prior knowledge measure can be found in Appendix F. Due to lack of power, this measure was not used in any analysis.

Argument Construction

Arguments were created using 16 different themes (e.g. death penalty, marijuana, etc…). Each theme was used to create four arguments: warranted pro, unwarranted pro, warranted con, and unwarranted con. For example the theme of gun laws may be made into the following arguments.

Warranted pro: The U.S. should decrease gun laws because it is our Second Amendment right to bear arms.

Unwarranted pro: The U.S. should decrease gun laws because many people already own them

Warranted con: The U.S. should not decrease gun laws because school shootings are becoming more frequent.

Unwarranted con: The U.S. should not decrease gun laws because they are becoming more expensive.

This design resulted in 64 total arguments to be evaluated with four arguments constructed for each of 16 topics. Half of the 16 topics were controversial (gun control, abortion, etc.) and the other half were neutral (speed limits, ATM fees etc. Reasons for the arguments were constructed to meet a few criteria. First, each reason was made to be a true statement. Second, the reasons used referents (i.e. “they”, “it”) in order to reduce semantic overlap when referring to an object in the claim. And third, the reasons were constructed to be similar in character length. A complete list of full arguments is presented in Appendix C.

Evaluation Task

Participants were asked to evaluate the logical quality of each of the 64 arguments. The task asked participants to rate the quality of the arguments on an 8 point Likert scale ranging from 1(very low quality) to 8(very high quality). This is similar to the task given in Experiment 1. The instructions for this task are given in Appendix G.

Control Task

Participants in the control task were given a probe recognition task. They were instructed to read each argument naturally and then to decide whether a probe word presented on the screen was present or not in that argument. Participants in the control task read the same arguments as participants in the evaluation condition, the only difference being their purpose for reading.

Instructions for this task is found in Appendix H.

Reflective Thought Listing

This task asked participants to state what they were thinking while reading 8 arguments which they previously read. Edwards and Smith (1996) used a think-aloud in their experiment as well. The think-aloud occurred after participants had completed all of their argument strength judgments. During the think-aloud, participants were shown the claim of each topic and asked to list arguments that came to mind as they considered each claim. One issue with this task is that it does not ask participants to consider the reasons given in support of the claims they read. Participants are simply asked list what they think about the claims themselves. In order to better understand what participants think while they process and evaluate arguments, I piloted a slightly different thought listing task. For this task I retrospectively asked participants to type what they were thinking while reading a pseudo-random selection of 8 arguments which they saw before. This pseudo-random selection ensured that participants received an equal spread of arguments which were pro, con, warranted, unwarranted, and as well as varying strengths of attitudes about. These 8 arguments included 1 argument related to each of the 8 controversial themes. This should be a better measure of what participants were thinking during reading of the whole argument and allowed participants to state how the reasons and implicit warrants in the arguments may have played a role in their processing. The instructions for this task were:

> "For the next part of the study, we are interested in your thoughts about some of the arguments you were asked to read.
>
> You will only have to answer this for 8 of the arguments you read. Please try to think back to what you were thinking when you originally read each argument. When you have typed your thoughts for each argument, press Next to move to the next argument."

"What were you thinking when you read the following argument?"

Procedure

The pretest survey was given 10-14 days prior to completion of part 2. It included attitude strength ratings for the 16 topics, the attitude certainty ratings for the 16 topics, and the prior knowledge ratings for the 16 topics. This survey was presented on Qualtrics through Mechanical Turk. Participants self-selected into the study through Mechanical Turk. Participants were paid $0.50 for completion of part 1 and $2.50 for completion of part 2.

Upon completion of the pretest survey, participants were invited back to complete part 2. All participants who participated in the pretest survey were invited to complete the second portion of the experiment. Participants were be randomly assigned to either be in the evaluation task condition or the control condition. This portion of the experiment was again presented on Qualtrics through Mechanical Turk. Responses and timing data was collected for all measures given. After the informed consent process, participants were given general instructions about the tasks they were going to be asked to complete. Following instructions, participants were given 4 practice trials to become familiar with the procedure. After practice, participants completed the main task. Participants in the evaluation condition were first shown a claim, then upon a spacebar press, the reason associated with the claim. Followed by a screen prompting their evaluation. Participants in the control condition completed the same procedure as those in the evaluation task condition with the exception of completing the evaluation task. Instead, participants were shown a word and prompted to decide whether that word was present in the argument they had just read. Completion of the experiment was self-paced. Sixty-four items were presented to each participant. Reading times were collected during the reading of the claim and reason. The reading time of the reason is the main reading time dependent variable, since

that is when warrant creation and evaluation is likely to occur. After participants completed reading and responding to all 64 arguments, they were asked to complete the thought listing task. After completing the evaluation and thought listing tasks participants completed a demographics survey. Figure 6 shows the timeline of the main task procedure.

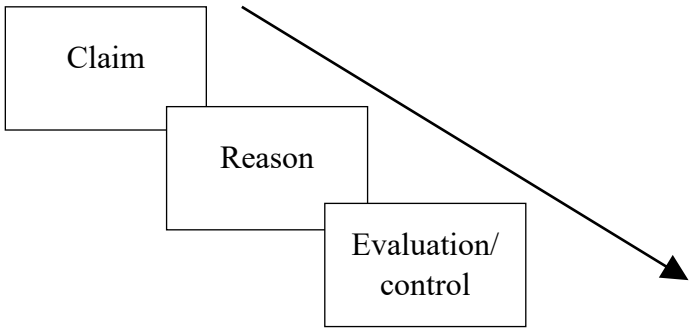

Figure 6. Task order during the main study.

Data Transformation

Pro and Con arguments were categorized as attitude consistent or attitude inconsistent using a similar procedure as in Experiment 1. This can be seen in Table 1.

Additionally, reason reading time data was transformed to time-per-character by dividing by the number of characters for each reason. This was done to control for reason character length across arguments.

Scoring Thought Listing Protocols

Thought Listing protocols were scored using a system based on Edwards and Smith (1996) as well as Taber and Lodge (2006). These protocols were scored based on overall positive and negatives statements made about features of the arguments and additions to the arguments. Participants could make positive or negative evaluations about the following four argument features: claim, reason, warrant, or additional statements. For this analysis, I only considered the

presence or absence of these evaluations, not a count of how many they made within each feature category. This allowed argument positivity scores to range from -4 to 4 for each argument. Scores were then transformed for attitude consistency, using similar methods as the evaluation transformation. Finally, scores were averaged across attitude consistency and argument quality. Inter-rater reliability was achieved with an independent scorer ($k = .91$).

Sample responses are shown in Table 2.

Table 2.

Sample thought listing protocols of each type for a warranted and unwarranted argument. The part in italics was not part of the response scored in for the type identified.

Warranted argument: The U.S. should decrease gun laws because it is our Second Amendment right to bear arms.

	Claim	**Reason**	**Warrant**	**Additional statements**
Positive:	I have to agree with that.	That the right to bear arms is a fact *so it supports the argument to decrease gun laws.*	*That the right to bear arms is a fact* so it supports the argument to decrease gun laws.	*On the surface that's true,* but where is the cutoff point? Should a person be allowed to own AK-47s? Bazookas? Rocket launchers? Should anyone be able to purchase a firearm without a background check or waiting period? There are many nuances and things that should be just plain common sense in any modern-day argument which supports Second Amendment rights.
Negative:	I disagree with this statement. *I agree with the Second Amendment but also think that we need increased gun laws.*	N/A	N/A	Decreasing gun laws also decreases childrens' rights not to get shot when they are in school.

Unwarranted argument: Governments should ban pornography because it is free on many sites on the internet.

	Claim	**Reason**	**Warrant**	**Additional statements**
Positive:	Pornography is wrong and an offense to women. *It encourages violence toward women and disrespect. It's availability should be legislated.*	The fact that it is free *should have no bearing on whether it is illegal or not.*	N/A	*Pornography is wrong and an offense to women*. It encourages violence toward women and disrespect. It's availability should be legislated.
Negative:	I disagree with this statement.	N/A	I was basically thinking that this statement makes no logical sense to me. Many things are free - things shouldn't be banned just because they are free.	So? It doesn't harm one soul

Results

Argument Evaluation

A 2 Quality (warranted vs. unwarranted) x 2 Attitude Consistency (consistent vs. inconsistent) x 4 Attitude Strength (strongly disagree vs. weakly disagree vs. weakly agree vs. strongly agree) within-participants repeated measures ANOVA was conducted on argument evaluation ratings. Attitude strength is based on participants' initial attitude agreement judgements.

All three main effects were significant. Participants rated warranted arguments (M=5.88, SD=0.99) higher than unwarranted arguments (M=2.42, SD=0.68) ($F(1,20) = 200.85$, $p < .001$, $\eta_p^2 = .91$). Participants rated attitude consistent arguments (M=4.37, SD=0.22) higher than attitude inconsistent arguments (M=3.93, SD=0.23) ($F(1,20) = 13.17$, $p = .002$, $\eta_p^2 = .40$). The main effect of agreement was also significant ($F(1,20) = 2.82$, $p = .047$, $\eta_p^2 = .12$). Mauchly's Test of Sphericity was also non-significant $\chi2(5) = 7.14$, $p = .211$, so sphericity is assumed. There was a significant consistency x agreement interaction $F(2.033, 40.66) = 15.36$, $p < .001$, $\eta_p^2 = .43$). Sphericity cannot be assumed ($\chi2(5) = 13.15$, $p = .022$, so the Greenhouse-Geisser adjustment was used. Figure 7 shows this effect. The Consistency by Quality interaction found in Experiment 1 was not replicated, $F(1, 20) = .034$, $p = .855$. No other significant effects were found. Table 3 shows the average quality evaluations for the attitude consistency by quality by attitude strength interaction.

The significant main effect of quality provides support for the Logical Dominance Hypothesis. While making evaluations, participants were able to recognize and rate unwarranted arguments lower than warranted. Additionally, the attitude consistency effect only occurred for warranted items, meaning that in order for attitudes to impact evaluations, the arguments had to

meet at least a minimal bar of quality. The consistency by agreement interaction also shows partial support for both the Disconfirmation and Bolstering and Defending hypotheses. Attitude strength seems to matter to the attitude consistency effect, but not necessarily where expected.

Table 3

Average Quality Ratings (with Standard Deviation) for Attitude Consistency x Quality x Agreement conditions in Experiment 2.

Agreement	Quality	Attitude Consistent	Attitude Inconsistent	Attitude Consistency Effect
Strongly Disagree	Warranted	5.37 (1.54)	5.9 (1.37)	-0.5
	Unwarranted	2.25 (1.09)	2.38 (1.23)	-0.1
Weakly Disagree	Warranted	5.23 (1.95)	6.16 (1.98)	-0.9
	Unwarranted	2.05 (1.3)	2.64 (1.82)	-0.6
Weakly Agree	Warranted	7.06 (0.55)	5.16 (1.58)	1.9
	Unwarranted	2.93 (1.29)	1.71 (0.83)	1.2
Strongly Agree	Warranted	6.66 (1.21)	5.44 (1.54)	1.2
	Unwarranted	3.37 (1.46)	2.06 (0.91)	1.3

Note. The final column shows the average Attitude Consistency Effect as a difference score for the Attitude Consistent rating minus the Attitude Inconsistent rating.

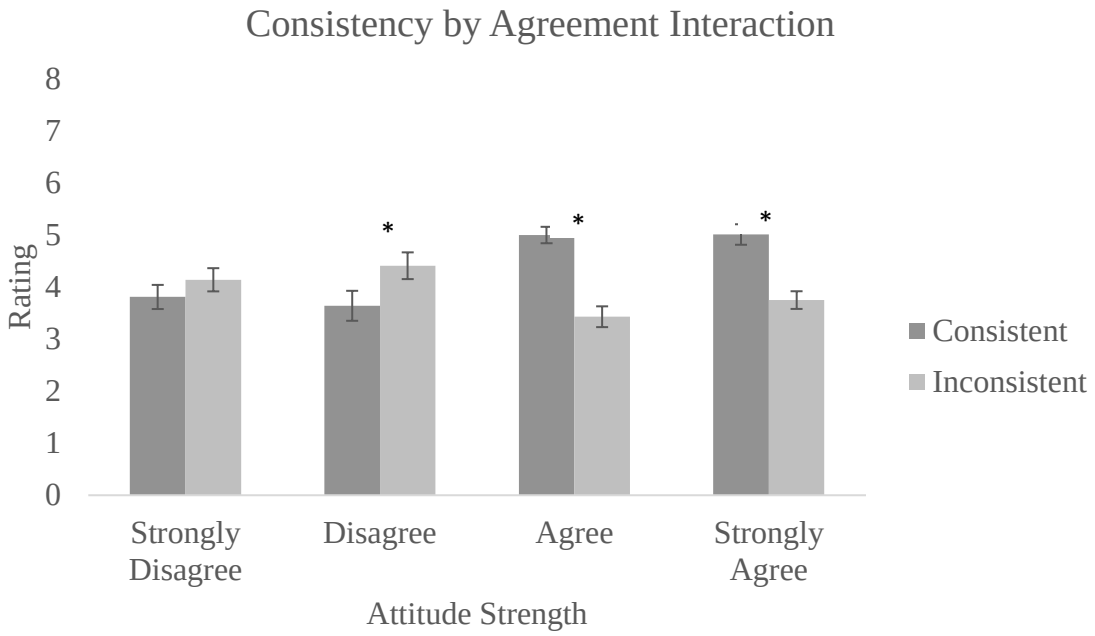

Figure 7. Consistency by Attitude Strength interaction for argument evaluations.

* indicates significant at the $p < .05$ level.

Reason Reading Time

Effect of Task. A 2 Quality (warranted vs. unwarranted) x 2 Attitude Consistency (consistent vs. inconsistent) x 4 Attitude Strength (strongly disagree vs. weakly disagree vs. weakly agree vs. strongly agree) x 2 Task (evaluation task vs. control) mixed effects ANOVA was conducted on reason reading time. Table 4 shows the average reason reading times for the Attitude Consistency by Quality by Attitude Strength. Attitude strength is based on participants' initial attitude agreement judgements.

Table 4

Average Reason Reading Time per Character in milliseconds (with Standard Deviation) for Attitude Consistency X Quality X Agreement conditions in Experiment 2.

		Warranted		Unwarranted	
		Consistent	Inconsistent	Consistent	Inconsistent
Evaluative task	S. Disagree	92 (9)	78 (7)	87 (8)	79 (9)
	W. Disagree	87 (6)	67 (5)	87 (7)	98 (10)
	W. Agree	68 (6)	87 (6)	79 (7)	77 (6)
	S. Agree	69 (6)	90 (9)	81 (7)	78 (6)
Control task	S. Disagree	48 (9)	45 (7)	49 (8)	44 (9)
	W. Disagree	43 (6)	39 (5)	45 (7)	43 (10)
	W. Agree	45 (6)	42 (6)	49 (7)	47 (6)
	S. Agree	42 (6)	46 (9)	51 (7)	47 (6)

Note. The table shows the means for both the evaluative task and control task.

There was a significant effect of task on reason reading time (ms), $F(1,40) = 19.67$, $p < .001$, $\eta_p^2 = .33$) Participants in the experimental condition took longer to read the reasons (M = 82, SD = 27) than those in the control condition (M = 45, SD = 27). There was also a marginally significant 4-way quality by attitude consistency by agreement by task $F(2.48, 99.07) = 2.68$, $p = .062$, $\eta_p^2 = .063$. Sphericity cannot be assumed ($\chi 2(5) = 15.61$, $p = .008$, so the Greenhouse-Geisser adjustment was used. Because of this difference, a separate ANOVA was conducted on reason reading times for only those in the evaluation condition.

The pattern of results for task provide support for the Evaluative Mindset hypothesis. Any attitude consistency effect was reliant on participants being in the evaluative task and therefore in an evaluative mindset.

Effect of Evaluation Task. A 2 Quality (warranted vs. unwarranted) x 2 Attitude Consistency (consistent vs. inconsistent) x 4 Attitude Strength (strongly disagree vs. weakly disagree vs. weakly agree vs. strongly agree) within-participants repeated measures ANOVA was conducted on reason reading time for only those in the evaluation condition. Attitude strength is based on participants' initial attitude agreement judgements.

There were no significant main effects on reason reading time. The main effect of quality was non-significant ($F(1,20) = 1.90$, $p = .184$). The main effect of attitude consistency was non-significant ($F(1,20) = .02$, $p = .880$). The main effect of agreement was non-significant ($F(1,20) = 1.61$, $p = .198$). There was a marginally significant consistency by agreement interaction on reason reading time $F(1.916, 38.319) = 3.026$, $p = .062$, $\eta_p^2 = .13$), shown in Figure 8. Sphericity cannot be assumed ($\chi2(5) = 17.24$, $p = .004$, so the Greenhouse-Geisser adjustment was used. There was also a significant higher-level interaction between consistency, quality, and agreement ($F(2.29, 45.87) = 3.65$, $p = .028$, $\eta_p^2 = .15$), shown in Figures 9 – 12. Sphericity cannot be assumed ($\chi2(5) = 11.25$, $p = .037$, so the Greenhouse-Geisser adjustment was used.

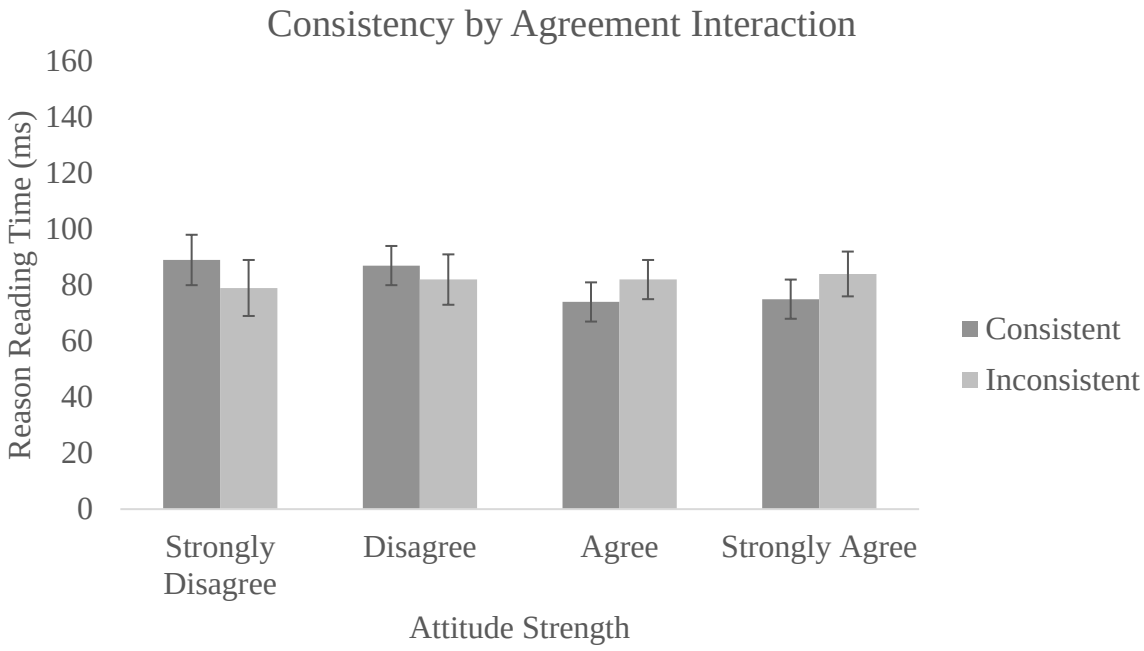

Figure 8. Consistency by Attitude Strength on reason reading time.

Figure 9. Consistency by Quality for Strongly Disagree items on reason reading time.

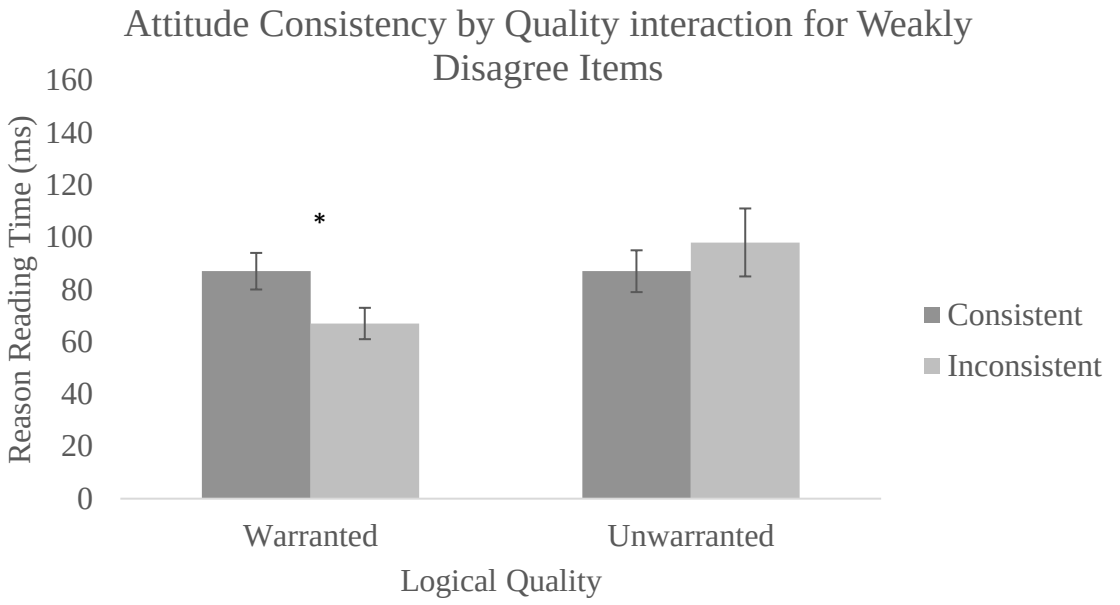

Figure 10. Consistency by Quality for Weakly Disagree items on reason reading time.

* indicates significant at the $p < .05$ level.

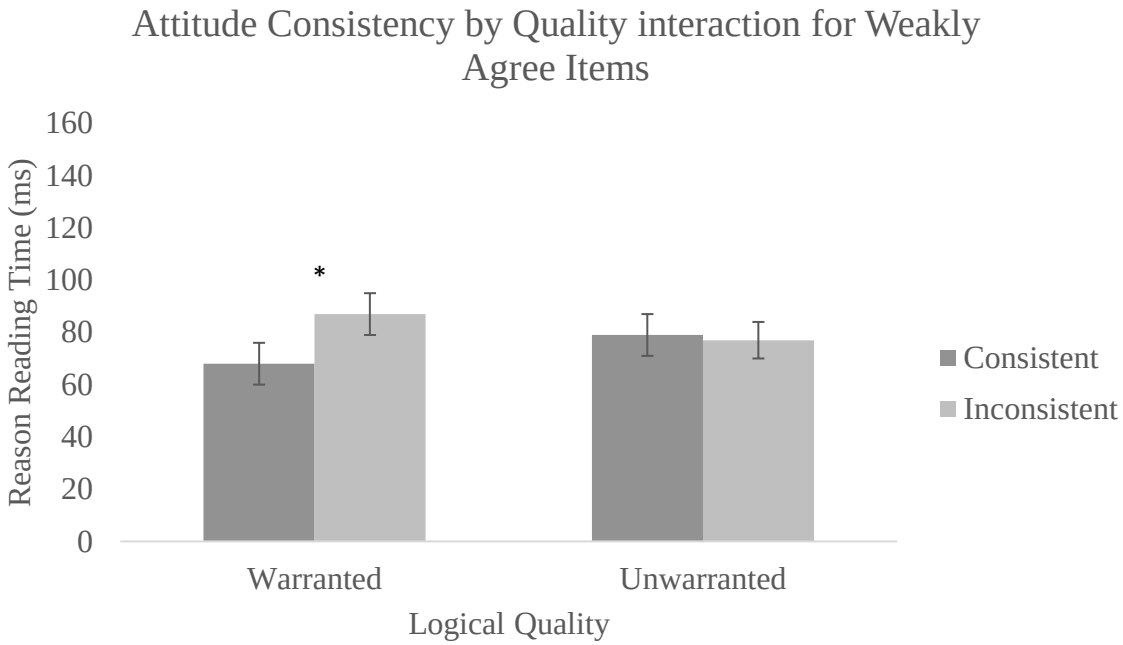

Figure 11. Consistency by Quality for Weakly Agree items on reason reading time.

* indicates significant at the $p < .05$ level.

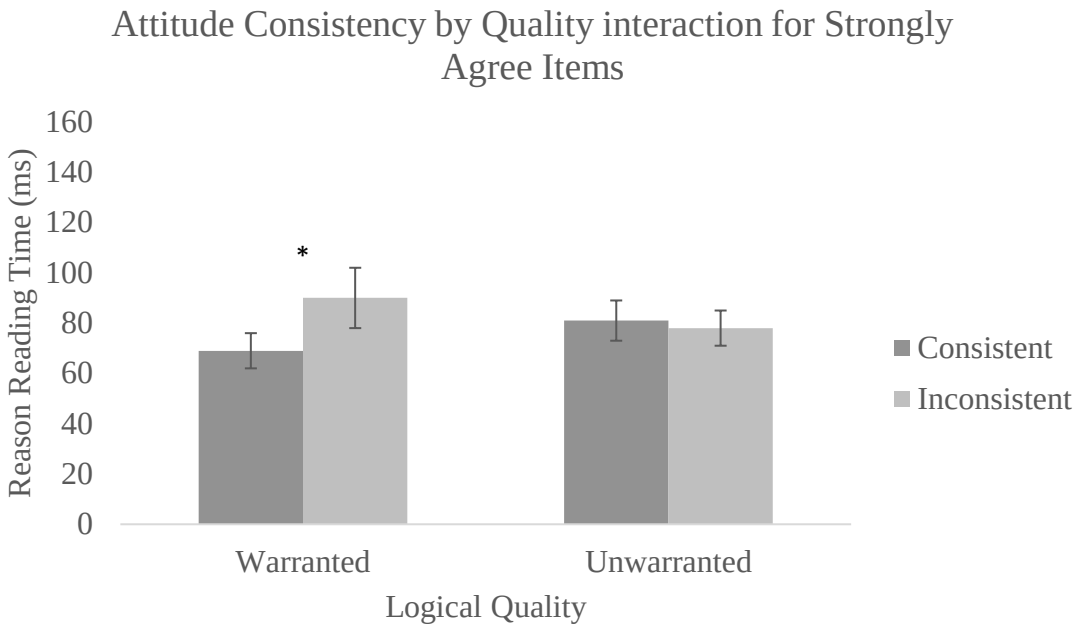

Figure 12. Consistency by Quality for Strongly Agree items on reason reading time.

* indicates significant at the $p < .05$ level.

The results on reason reading time do provide some partial support for both the Disconfirmation and Bolstering and Defending hypotheses. Differences in reading time were related to different levels of attitude strength, though not necessarily where would have been expected.

Thought Listing Positivity Scores

A 2 (warranted vs. unwarranted) x 2 (attitude consistent vs. attitude inconsistent) within-participants repeated measures ANOVA was conducted on positivity scores. There was a main effect of quality on positivity scores, $F(1,19) = 36.03$, $p < .001$, $\eta_p^2 = .66$. Participants reported more positive thoughts about possible argument features for the warranted arguments (M = .90, SD = .78) compared to the unwarranted arguments (M = -.55, SD = .71). The effect of attitude consistency was not significant, nor was the interaction.

Discussion

Experiment 2 addressed three factors that could potentially set boundary conditions on the influence of prior attitudes on argument processing. In this experiment, each of these factors were found to set a limit on the effect in terms of quality ratings and/or reason reading time. Overall, there was an effect of attitudes on both ratings and reason reading times, both in the expected direction.

The first research question examined whether the attitude consistency effect occurred for argument evaluation even if arguments were unwarranted (the reasons did not even weakly support the claim). The results of Experiment 2 show that overall, this population of participants (M-Turk) was sensitive to argument quality. This is a new finding. In fact, the average difference between warranted and unwarranted arguments was actually larger for Experiment 2 (M difference = 3.46) than Experiment 1 (M difference = 2.46). It should also be noted that this effect was actually larger than the effect of attitude consistency. More importantly, however, to address the first research question, the critical question is whether these two factors interact. On the one hand, after reading a counter-attitudinal claim, the reader may increase their scrutiny for the upcoming reason. If so, it was expected that the attitude consistent arguments would be rated high than inconsistent arguments, regardless of the quality of the argument. If this was due to increased scrutiny, then there should also be longer reason times for both types of arguments when the claims are attitude inconsistent then when they are attitude consistent. On the other hand, if like Experiment 1, the scrutiny is only triggered when the argument rises to a certain level of "goodness", then there should be an interaction of argument quality and attitude consistency for both ratings and reason reading time. The results were that the interaction of argument quality and attitude consistency was not significant for ratings. This was not even close

to a trend. So unlike Experiment 1, the attitude consistency effect holds for both levels equally. It is unclear whether these differences in results were due to the population (undergraduates vs. Mechanical Turk workers), the testing situation (in person in a lab vs. online with no control of situation), motivation (class credit vs. money), or selection (only half of those who completed part 1 completed part 2 in Experiment 2). However, for both experiments, evaluation of the unwarranted arguments was not influenced by consistency with prior attitudes. Thus, when arguments were unwarranted, participants generally did not seem to give the argument the benefit of the doubt if they agreed. Participants were able to recognize that the quality of the argument was the focus of their evaluation when an argument was obviously "bad". However for Experiment 2, there was a significant Attitude consistency by Quality by Agreement on reason reading time. I will discuss this in more detail for research question 2, the important point here is that all of the effects of attitude consistency were found only on the reading time for the reasons for the warranted arguments. Thus, although there were no effects on rating for Experiment 2, there were effects on the processing time of the reason differentially depending on whether the arguments were minimally warranted. The effects of bolstering and scrutiny are limited to warranted items.

The second research question examined whether the attitude consistency effect occurred on ratings and reason processing time at all levels of attitude strength. The key hypothesis was that it would be either a consistent effect for all levels of attitude strength (as suggested by the correlation found in Edwards & Smith, 1996) or stronger for arguments that the participant has strong attitudes about regardless of direction (which was actually tested by Edwards & Smith, 1996 and supported by Voss et al., 1993). I found a significant Agreement by Attitude Consistency interaction on ratings. Attitude consistent arguments were rated higher when

participants initially agreed with the claim (either weakly or strongly), but the reverse was true for claims that were initially weakly disagreed with. Interestingly, there was no effect of attitude for arguments when participants strongly disagreed with the claim. This pattern replicated the findings from Experiment 1.

The findings for reason times also support the conclusion that the attitude consistency affect is not consistent across all levels of strength. The Attitude consistency by Quality by Attitude Strength interaction was significant for reason reading times. When reading arguments that the participant initially agreed with, even weakly, they spent longer reading attitude inconsistent reasons than attitude consistent reasons but only for a warranted claims. This finding is consistent with the basic scrutiny effect. But the scrutiny effect does not hold when those arguments are unwarranted. At least when evaluating arguments for quality, the reason has to be minimally relevant to lead to longer processing or more scrutiny. Now for those arguments with a claim that the participant initially disagreed strongly with, there was no extra time to read the reason based on consistency with their attitude. Finally, for those arguments with claims that the participant initially weakly disagree with, there is a reverse attitude consistency effect. They spent longer reading attitude *consistent* reasons than attitude inconsistent reasons. This is consistent with a bolstering effect. Again, this longer reading time only occurred when the reasons were minimally supporting of the claim (i.e., warranted arguments). Overall, these findings are not consistent with those of Edwards and Smith (1996) that the pattern would be different for agreeing vs disagreeing. I will elaborate on this more in the general discussion.

The final research question examined whether the attitude consistency effect occurred only for an evaluative task. The findings support the conclusion that an evaluated task is needed. There were no attitude consistency effects for the control task that only required reading, but not

evaluation of the arguments. This lack of effect on attitudes for a non-evaluative task could still be consistent with Voss et al., (1993) finding of an effect of very fast activation of attitudes and values even when reading only for comprehension. It is possible that in the probe task, attitudes were activated and available to the participants, but they were not used while completing the task. Thus, one's attitudes could have an effect on processing of the reason, but they do not seem to, at least not during a superficial probe task.

For the thought listing responses, no attitude consistency effect was found. Neither the main effect of attitude consistency nor the interaction between argument quality and attitude consistency was found. This cursory analysis did not replicate the disconfirmation effect found by Edwards and Smith (1996). This means I did not find support for scrutiny in the thought listing analysis. However, the lack of an attitude consistency effect was not necessarily due to a floor effect given that there was a significant effect of argument quality. It is possible that the consistency effect could have been countered by a bolstering effect, which Edwards and Smith (1996) would not account for. A third possibility is that the presentation of the entire argument, instead of just the claim, and doing a slightly deeper scoring of the protocols lead to not replicating the Edwards and Smith (1996) findings. Though, there are also several weaknesses to this task. Participants only completed this task for 8 arguments and all participants responded to the same 8 arguments. Also, given the quasi-random assignment of attitude consistency, there was an unequal number of responses for participants across arguments related to which were or were not attitude consistent with that given participant. Another possibility is that participants became less engaged in the task, as it was the last task completed. While most responses seemed to show compliance, it is possible that participants responded differently than they would have if

the task was done earlier in the experiment. While there are issues with the task, I believe it shows promise for a more complete set of categories to be used in a thought listing analysis.

These results show, across two very different situations and populations, that attitudes can influence argument evaluation and processing. Experiment 2 also suggests a role for both bolstering and scrutiny in argument evaluation. Finally, they show that an individual difference factor, a material factor, and a task factor all can provide limits to the general attitude consistency effect.

CHAPTER 4: GENERAL DISCUSSION

Over the past 45 years, many have found an attitude consistency effect (Clark & Wegener, 2013, Edwards & Smith, 1996; Lord, Ross, & Lepper, 1976; Wolfe & Kurby, 2017). This dissertation was able to find an attitude consistency as well, using short (single sentence) informal policy arguments about a variety of topics. Additionally, I tested three boundary conditions to the effect.

In this dissertation, I have tested three very different sources of potential boundaries on the attitude consistency effect which were previously unexplored: an individual difference factor of the strength of one's attitude for a particular topic, a materials factor of clearly unwarranted arguments, and a task factor of a non-evaluative task. Overall, I found a rather robust attitude consistency effect for these short claim-reason arguments on ratings (Experiments 1 and 2) and in reason reading times (Experiment 2). However, I also found this effect to be qualified in important ways by each of these factors. The most consistent finding was an interaction of Attitude consistency by Quality by Attitude strength in ratings and in reason reading time. In terms of ratings, for both studies there was the expected effect of attitudes on arguments when the claims were in the form that they agreed with or strongly agreed with. Participants rated the arguments higher that were attitude consistent than when attitude inconsistent. There was no effect of attitude on arguments for claims participants strongly disagreed with. In the case where the participant was given a claim they weakly disagreed with, they rated it as stronger when it was opposite of the original perspective (but only for Experiment 2). For reason reading time (Experiment 2 only), a similar pattern was found. There was a consistent effect of attitudes when

reading reasons for claims they strongly and weakly agreed with but the reverse was found for claims they had rated as weakly disagree with and no effect for those they strongly disagreed with. These findings provide partial support for the Disconfirmation Hypothesis as well as the Bolstering and Defending Hypothesis. All of this was only found for the evaluative tasks (strength and logical quality ratings) but not for a control task. This finding supports the Evaluative Mindset Hypothesis. Thus, each of the boundary factors had some impact on the effects of prior attitudes on argument processing. In the next section, I will review each research question separately.

This dissertation was able to identify some boundary conditions to the Attitude Consistency Effect. But how do these findings relate to prior research? Research from a cognitive perspective found that even with training, people are only around 80% accurate at evaluating the logical quality of arguments (Larson et al., 2009). The implication of the results of this dissertation suggest that attitudes do play some role in this evaluation process, especially with warranted arguments. Though, this effect is much less robust than the effect of the quality of the arguments themselves. Overall, while reading arguments, people are still able to relatively accurately represent what they read, and make their evaluations based on the content of the argument, with little impact of their attitudes. This result shows partial support for the Logical Dominance Hypothesis. While people are able to reliably evaluate the logical quality of arguments, and this effect is very strong, participants' attitudes still did have some impact on evaluations.

Similarly, research on argument evaluation from a social psychology perspective has used at least minimally acceptable arguments, while sometimes varying the degree of acceptability from strong to weak (Clark et al., 2008), but never to the point of providing a

logically unacceptable argument. The results from this dissertation suggest that this methodology may or may not be warranted. In Experiment 1, attitudes did not affect the ratings of logically unacceptable arguments, but Experiment 2 did not find this interaction, attitudes impacted both logically acceptable and unacceptable arguments equally. But, these unacceptable arguments were always evaluated as having lower quality than acceptable arguments. Ultimately, people are still able to identify "bad" arguments, regardless of their attitudes.

I was also able to find support for some level of scrutiny, as well as bolstering. Participants scrutinized only warranted arguments that they agreed or strongly agreed with (Experiment 2). Participants also engaged in bolstering for warranted arguments that they disagreed with (but not strongly disagreed with). This effect of additional processing to scrutinize arguments or bolster one's own side was limited to a 3 way interaction. Within this interaction there were 8 opportunities to either scrutinize or bolster but these only happened 3 times, with two instances of scrutiny and one instance of bolstering. Of those three instances, none occurred when arguments were unwarranted, which may suggest that people are able to identify unwarranted arguments and dismiss them without excessive processing, regardless of their attitudes. Another possibility is that when reading unwarranted attitude consistent argument, people may try to make extended inferences in order to "make sense" of the argument. More work should be done to investigate the type of inferencing that may be occurring during this additional processing time.

Overall these results show that the Disconfirmation Model (Figure 1) proposed by Edwards and Smith (1996) may need to be modified. These results question the strong claim that every time one encounters an attitude inconsistent argument, they will engage in an effortful memory search (scrutiny). Instead, the Discrepancy Motives Model proposed by Clark and

Wegener (2013) (Figure 2) was correct in the addition of bolstering, as people do engage in some effortful processing of arguments they agree with under some circumstances. However, more research needs to be done to explore more triggers for and boundaries to bolstering and scrutiny. I think the biggest modification to make to these models is the addition of more boundaries to when these effects occur. The results of this dissertation have shown that the materials used, strength of attitude, and task can impact when attitudes may play a role on evaluations. These types of boundaries should be added to form a more complete model of the effect.

Although this dissertation focused on attitudes and argument processing, it fits in a larger research program. Several researchers have found that one's prior beliefs can influence the evaluation on arguments across multiple documents (McCrudden & Barnes, 2016), information selection (confirmation bias, Nickerson, 1998), text comprehension (text-belief consistency effect; e.g., Eagly & Chaiken, 1993; Maier & Richter, 2013), and lack of updating (continued misinformation effect; e.g., Chinn & Brewer, 1993; Johnson & Seifert, 1994; Ross, Lepper, & Hubbard, 1975), and argument production (Myside Bias; e.g. Wolfe, Britt, & Butler, 2009). While I discussed the difference between attitudes and beliefs earlier in the paper, I think it is worth revisiting here. According to Wolfe and Griffin (2017), beliefs are a "position on a question of fact…" (Wolfe & Griffin, pg. 296), while attitudes are "a valence affective preference for something…" (Wolfe & Griffin, pg. 296). This distinction is important as people may scrutinize and bolster differently when reading arguments about beliefs (whether something is true) compared to attitudes (whether we should or should not do something). For example, a claim used in this study, "We should decrease gun laws" would elicit an attitude about ones preference for decreasing gun laws or not. Whereas an argument such as "gun laws in the U.S are strict" would elicit one's beliefs about the truth value of the statement. This distinction, and

the results of this dissertation focusing on attitudes, opens the door for further research in order to investigate how beliefs and attitudes may function differently in the belief bias research mentioned above. This is important, especially in the context of belief bias in science. Being able to ignore ones attitudes about scientific findings and focus on arguments presented should lead to more consistency and less belief bias within the scientific community.

Ultimately our attitudes about experiences and topics certainly play a role in how we navigate our world. Attitudes are always present (Voss et al., 1996) but they doesn't mean they will always impact our behaviors. They still may impact our processing of situations, being able to identify that our attitudes impact our processing and then correct for that bias in our behavior is extremely important. This is particularly important in the current world of social media when echo chambers are so prevalent. People often encounter, and even seek out, information and arguments that are in alignment with their own attitudes. Some of these arguments can be extremely low quality and even unwarranted, but many people still find them convincing and will repost them, spreading the bad argument to others. In this situation, you may not be in an evaluative mindset and therefore it may be that attitudes would lead to greater acceptance of unwarranted arguments on social media. Helping people become better at recognizing how their attitudes impact how they see this type of information can be a step in the right direction.

Limitations

While the results of this dissertation are promising, there are several limitations to the current work. Firstly, sample sizes were relatively small in both Experiment 1 and 2. While Experiment 1 did have more participants, these were split across two conditions for the analyses. Secondly, data in Experiment 2 was collected online using Mechanical Turk. Not only does this change the population between Experiment 1 and 2, it also raises issues related to collecting

timing data online. While there hasn't been a lot of research on the collection of timing data online, some has suggested that the data can be valid, at least for detecting differences as small as 20ms (Simcox & Fiez, 2014). There is also a potential selection bias, as participants had to actively choose to come back and participate in Experiment 2, which about half did not. The task itself used arguments which were fairly short, simple 1 sentence per argument. This type of informal argument does not necessarily represent all of the arguments one may encounter in daily life. Additionally, for the pre-test measurement of attitudes all of the arguments were presented in their "pro" form (we should do X). This meant that every "con" argument encountered later in the study contained the word "not". This imbalance could have impacted the results of Experiment 2.

Future Directions

One interesting area to explore is the use of longer arguments while investigating the impact that inducing scrutiny or bolstering early on in argument processing may have on later processing. If someone is reading an article and first encounters an attitude inconsistent argument, they will likely engage in scrutiny. How does that impact how they processing attitude consistent or inconsistent arguments later in the articles? How does that pattern change if the first encounter an argument and engage in bolstering instead? These types of questions may better generalize to how people tend to engage with information in their daily lives. Along these lines, people often encounter arguments as part of a larger situation. Their highest level goal may not be simply to evaluate arguments, instead, evaluating arguments may be something they have to do to meet a larger goal. For example, in a courtroom. Someone on a jury must hear and evaluate many arguments in order to come to their conclusion. In conjunction with other evidence which

may help build ones' attitude, the effect of attitude consistency and giving the "author" the benefit of the doubt if they give a bad or less warranted argument may be more prominent.

I think exploring the distinction between attitudes and beliefs, specifically related to how they impact processing differently would be a worthwhile direction as well. Understanding how holding a belief about a fact differs from holding an attitude about ones affective preference for or against that fact could have implications on framing in persuasive arguments.

Conclusions

This dissertation contributes to the existing literature by finding evidence for boundary conditions to an attitude consistency effect. Firstly, the logical quality of the arguments matters. While research before has looked at the strength of arguments (Clark et al., 2008) and their impact on belief consistency effects, only one has used completely unwarranted arguments (Wolfe & Kurby, 2017), but they did measured argument evaluation using a potentially insensitive binary measure. Second, I found that attitude direction and strength are important to this effect as well. Strongly agreeing with a claim seems to be fundamentally different than strongly disagreeing with it, even when reading arguments that are consistent with your disagreement or agreement. Attitude consistency alone doesn't account for differences in reading times. A third, task matters to the effect. Voss et al. (1993) found that attitudes and values are activated for strongly held beliefs can happen as quickly as comprehension occurs. My experiment found that while those attitudes may be activated, they may only impact processing if the task requires their use.

REFERENCES

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1061–1069.

Blair, J. A., & Johnson, R. H. (1987). Argumentation as dialectical. *Argumentation*, 1(1), 41-56.

Bohn-Gettler, C. M., & McCrudden, M. T. (2018). Effects of task relevance instructions and topic beliefs on reading processes and memory. *Discourse Processes, 55*, 410-431.

Britt, M. A., Kopp, K., Durik, A.M., Blaum, D., & Hastings, P. (2016). Identifying general cognitive abilities involved in argument comprehension and evaluation. *Zeitschrift für Pädagogische Psychologie, 30*, 79-95. DOI: 10.1024/1010-0652/a000173

Britt, M.A, & Larson, A. A. (2003). Constructing representations of arguments. *Journal of Memory and Language, 48*, 794-810. doi: 10.1016/S0749-596X(03)00002-0

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63,* 1–49.

Clark, J. K., Wegener, D. T., & Fabrigar, L. R. (2008). Attitudinal ambivalence and message-based persuasion: Motivated processing of proattitudinal information and avoidance of counterattitudinal information. *Personality and Social Psychology Bulletin, 34,* 565-577.

Clark, J. K., & Wegener, D. T. (2013). Message position, information processing, and persuasion: The discrepancy motives model. *In Advances In Experimental Social Psychology* (Vol. 47, pp. 189-232). Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Erlbaum.

Dandotkar, S. (2012). *Readers' Reliance on Semantic and Logical Relatedness when Evaluating Arguments*. Doctoral dissertation, Northern Illinois University.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. Journal of Personality and Social Psychology, 71, 5–24.

Gilbert, M. (1991). The enthymeme buster: A heuristic procedure for position exploration in dialogic dispute. *Informal Logic, 13,* 159-166.

Gilead, M., Sela, M., & Maril, A. (2019). That's my truth: Evidence for involuntary opinion confirmation. *Social Psychological and Personality Science, 10,* 393–401.

Hample, D. (2005). *Arguing: Exchanging Reasons Face to Face*. Mahwah, NJ: Erlbaum Publishers.

Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1420–1436.

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

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

Larson, A. A., Britt, M. A., & Kurby, C. A. (2009). Improving students' evaluation of informal arguments. *The Journal of Experimental Education, 77*(4), 339-365. doi:10.3200/JEXE.77.4.339-366

Long, D. L., & Lea, R. B. (2005). Have we been searching for meaning in all the wrong places? Defining the" search after meaning" principle in comprehension. *Discourse Processes, 39*, 279–298.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098-2109. doi: 10.1037/0022-3514.37.11.2098

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

Maier, J., & Richter, T. (2014). Fostering multiple text comprehension: How metacognitive strategies and motivation moderate the text-belief consistency effect. *Metacognition & Learning, 9*, 54–71.

Maier, J., & Richter, T. (2016). Effects of text-belief consistency and reading task on the strategic validation of multiple texts. *European Journal of Psychology of Education, 31*, 479–497.

Maier, J., Richter, T., & Britt, M. A. (2018). Cognitive processes underlying the text-belief consistency effect: An eye-movement study. *Applied Cognitive Psychology, 32,* 171-185.

McCrudden, M. T., & Barnes, A. (2016). Differences in student reasoning about belief-relevant arguments: A mixed methods study. *Metacognition and Learning*, *11*, 275-303.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*, 131–157.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175–220.

Petrocelli, J. V., Tormala, Z. L., & Rucker, D. D. (2007). Unpacking attitude certainty: Attitude clarity and attitude correctness. *Journal of Personality and Social Psychology, 92*(1), 30.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*, 880-892.

Shaw, V. F. (1996). The cognitive processes in informal reasoning. *Thinking & Reasoning, 2*(1), 51-80. doi:10.1080/135467896394564.

Simcox, T., & Fiez, J. A. (2014). Collecting response times using amazon mechanical turk and adobe flash. *Behavior Research Methods*, *46*(1), 95-111.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science, 50*(3), 755-769.

Toulmin, S. E. (1958). *The Uses of Argument*. Cambridge, MA: Cambridge University Press.

Voss, J. F., Fincher-Kiefer, R., Wiley, J., & Silfies, L. N. (1993). On the processing of arguments. *Argumentation, 7*(2), 165–181.

Wiley, J. (2005). A fair and balanced look at the news: What affects memory for controversial arguments? *Journal of Memory and Language, 53*, 95-109.

Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). Argumentation schema and the myside bias in written argumentation. *Written Communication, 26*, 183–209.

Wolfe, M. B., & Griffin, T. D. (2017). Beliefs and discourse processing. *In The Routledge Handbook of Discourse Processes* (pp. 295-314). Routledge.

Wolfe, M. B., & Kurby, C. A. (2017). Belief in the claim of an argument increases perceived argument soundness. *Discourse Processes, 54*(8), 599-617.

APPENDIX A

EXPERIMENT 1 ATTITUDE INSTRUCTIONS

There are no right or wrong answers to these questions. It is your personal beliefs that interest us. Use the scale below to answer the questions. **If you strongly disagree with a statement, fill in 1. If you strongly agree, select 10.** If you more or less agree with a statement, select a number in between.

APPENDIX B

EXPERIMENT 1 ARGUMENTS

Warranted Pro

Sex education should be required in schools because it can help reduce teen pregnancy.

Banks should charge ATM fees because the fees help pay for maintenance of the machines.

The speed limit should be reduced to 55 mph because car accidents would be less likely.

Vaccinations should be required for all children because they prevent children from contracting diseases.

Hunting for sport should be allowed because it helps control animal populations.

Research on animals should be banned because it can be harmful to the animals.

Illinois should require stricter car emissions standards because they help reduce our carbon footprint.

People should be allowed to file for bankruptcy because it can help people gain financial control of their lives.

TV news should sensationalize violence because it helps news stations get higher ratings.

Pornography should be banned because it often objectifies women.

Curfews should be placed on teenagers because they are most likely to commit crimes at night.

Marijuana should be legalized because it is less harmful than many legal drugs.

Recycling should be federally mandated because it will help reduce the size of our landfills.

American students should be required to learn a second language because being bilingual increases your chances of getting a job.

Professional athletes should be tested for steroids because they give those athletes an unfair advantage.

Unwarranted pro

Sex education should be required in schools because schools teach many subjects.

Banks should charge ATM fees because banks are financial institutions.

The speed limit should be reduced to 55 mph because many people drive on interstates.

Vaccinations should be required for all children because vaccinations are given in shots.

Hunting for sport should be allowed because hunting often happens in the woods.

Research on animals should be banned because research uses the scientific method.

Illinois should require stricter car emissions standards because Illinois is a state with larger distances between cities.

People should be allowed to file for bankruptcy because most adults have bank accounts.

TV news should sensationalize violence because news stations provide valuable information to people.

Pornography should be banned because it is free on the internet.

Curfews should be placed on teenagers because curfews are rules.

Marijuana should be legalized because marijuana is a plant that grows naturally.

Recycling should be federally mandated because recycling has become more efficient than it was years ago.

American students should be required to learn a second language because many languages have a similar structure.

Professional athletes should be tested for steroids because athletes are in good physical shape.

Warranted Con

Sex education should not be required in schools because parents may want to educate their children about sex.

Banks should not charge ATM fees because they are unfair to people who use them.

The speed limit should not be reduced to 55 mph because it would increase travel times.

Vaccinations should not be required for all children because some children cannot safely take vaccines.

Hunting for sport should not be allowed because it is unethical to kill for fun.

Research on animals should not be banned because it helps us save the lives of humans.

Illinois should not require stricter car emissions standards because many people cannot afford cars which would pass the standards.

People should not be allowed to file for bankruptcy because money they owe is never paid back.

TV news should not sensationalize violence because people will think violence is more common than it actually is.

Pornography should not be banned because it involves consensual sexual activity between adults.

Curfews should not be placed on teenagers because they should have independence.

Marijuana should not be legalized because it could have unknown harmful effects.

Recycling should not be federally mandated because it would be unfair to states that produce more goods that need to be recycled.

American students should not be required to learn a second language because knowing a second language is not a necessary life skill.

Professional athletes should not be tested for steroids because athletes using steroids make sports more exciting.

Unwarranted Con

Sex education should not be required in schools because humans are animals and all animals have sex.

Banks should not charge ATM fees because bank accounts gain interest.

The speed limit should not be reduced to 55 mph because cars are designed to go faster.

Vaccinations should not be required for all children because vaccinations are safely given at a doctor's office.

Hunting for sport should not be allowed because many people don't know how to hunt.

Research on animals should not be banned because research is done at universities.

Illinois should not require stricter car emissions standards because these tests happen at car dealerships.

People should not be allowed to file for bankruptcy because big companies file for bankruptcy.

TV news should not sensationalize violence because local channels show the news.

Pornography should not be banned because it exists in many forms.

Curfews should not be placed on teenagers because curfews are usually at night.

Marijuana should not be legalized because it can be grown in mass quantities as a crop.

Recycling should not be federally mandated because different types of materials can be recycled.

American students should not be required to learn a second language because there are many languages spoken in the world.

Professional athletes should not be tested for steroids because athletes train all year around.

APPENDIX C

EXPERIMENT 2 ARGUMENTS

Warranted Pro Controversial

The U.S. should build a border wall between Mexico and the U.S. because it will help reduce drugs entering the U.S.

The government should raise the minimum wage to $15 an hour because the current rate is not a livable income level.

The U.S. should ban late term abortions because it ends a potential human life.

The U.S. should decrease gun laws because it is our Second Amendment right to bear arms.

The U.S. should legalize marijuana because it is less harmful than many legal drugs.

Governments should ban pornography because it often leads to objectifying women.

Schools should require sex education because it can help reduce teen pregnancy.

TV news should exaggerate violence because it helps news stations get higher ratings.

Warranted Pro Neutral

Banks should charge ATM fees because they help pay for maintenance of the machines.

The government should reduce the highway speed limit to 55 mph because car accidents would be less likely.

Towns should place curfews on teenagers because they are more likely to commit crimes at night.

The U.S. should federally mandate recycling because it will help reduce the size of our landfills.

The U.S. should require that students learn a second language because it will increase your chances of getting a job.

The U.S. should ban hunting for sport because it can lead to the excessive death of animals.

Sports organizations should test professional athletes for steroids because they give those players an unfair advantage.

Governments should allow research on animals because it helps us save the lives of humans.

Warranted Con Controversial

The U.S. should not build a border wall between Mexico and the U.S. because it won't really reduce illegal immigration.

The government should not raise the minimum wage to $15 an hour because small businesses won't be able to stay open.

The U.S. should not ban late term abortions because they can save the lives of some women.

The U.S. should not decrease gun laws because school shootings are becoming more frequent.

The U.S. should not legalize marijuana because it could have unknown harmful effects.

Governments should not ban pornography because it involves consensual activity between adults.

Schools should not require sex education because it may violate religious values.

TV news should not exaggerate violence because people will misjudge how common it is.

Warranted Con Neutral

Banks should not charge ATM fees because they will lose loyal clients.

The government should not reduce the highway speed limit to 55 mph because it would increase travel times.

Towns should not place curfews on teenagers because it infringes on their rights.

The U.S. should not federally mandate recycling because it is too costly for many business and families.

The U.S. should not require that students learn a second language because it is not a necessary life skill.

The U.S. should not ban hunting for sport because it helps control animal populations.

Sports organizations should not test professional athletes for steroids because it is their choice what to put in their bodies.

Governments should not allow research on animals because they may not by treated humanely.

Unwarranted Pro Controversial

The U.S. should build a border wall between Mexico and the U.S. because the border is about 2,000 miles long.

The government should raise the minimum wage to $15 an hour because the government sets the lowest pay rate.

The U.S. should ban late term abortions because it is an elective surgical procedure.

The U.S. should decrease gun laws because many people already own them.

The U.S. should legalize marijuana because it is a plant that grows naturally.

Governments should ban pornography because it is free on many sites on the internet.

Schools should require sex education because schools teach many different subjects.

TV news should exaggerate violence because news stations provide valuable information.

Unwarranted Pro Neutral

Banks should charge ATM fees because banks are financial institutions.

The government should reduce the highway speed limit to 55 mph because many people drive on interstates.

Towns should place curfews on teenagers because they are rules or boundaries. The U.S. should federally mandate recycling because it has become more efficient than it used to be.

The U.S. should require that students learn a second language because many of them have a similar structure.

The U.S. should ban hunting for sport because many people don't know how to do it.

Sports organizations should test professional athletes for steroids because they are in good physical shape.

Governments should allow research on animals because they have to spend their lives in laboratories.

Unwarranted Con Controversial

The U.S. should not build a border wall between Mexico and the U.S. because only four states share the border.

The government should not raise the minimum wage to $15 an hour because it is paid to workers in services industries.

The U.S. should not ban late term abortions because it is a medical procedure done by professionals.

The U.S. should not decrease gun laws because firearms are becoming more expensive.

The U.S. should not legalize marijuana because it can be grown in mass quantities as a crop.

Governments should not ban pornography because it exists in many forms including the web.

Schools should not require sex education because it teaches students about reproduction.

TV news should not exaggerate violence because local channels show the news.

Unwarranted Con Neutral

Banks should not charge ATM fees because bank accounts gain interest each year.

The government should not reduce the highway speed limit to 55 mph because cars are designed to go faster.

Towns should not place curfews on teenagers because they usually occur at night or after dark.

The U.S. should not federally mandate recycling because different types of materials can be recycled.

The U.S. should not require that students learn a second language because there are many dialects spoken in the world.

The U.S. should not ban hunting for sport because it often takes place in the woods.

Sports organizations should not test professional athletes for steroids because they train all year around.

Governments should not allow research on animals because it is based on the scientific method.

APPENDIX D

EXPERIMENT 2 ATTITUDE INSTRUCTIONS

Now we are interested in what you think about some of the topics we have asked you about. There are no right or wrong answers to these questions. It is your personal attitudes that interest us. Use the scale below to answer the questions. If you strongly disagree with a statement, press 1. If you strongly agree, press 4. If you more or less agree with a statement, select a number in between.

APPENDIX E

ATTITUDE CERTAINTY SURVEY

Attitude Clarity

1. How certain are you that you know what your true attitude on TOPIC really is?

2. How certain are you that the attitude you expressed toward TOPIC really reflects your true thoughts and feelings?

3. To what extent is your true attitude toward TOPIC clear in your mind?

4. How certain are you that the attitude you just expressed toward TOPIC is really the attitude you have?

Attitude Correctness

1. How certain are you that your attitude toward TOPIC is the correct attitude to have?

2. To what extent do you think other people should have the same attitude as you on this issue?

3. How certain are you that of all the possible attitudes one might have toward TOPIC, your attitude reflects the right way to think and feel about the issue?

APPENDIX F

PRIOR KNOWLEDGE SURVEY

Please indicate how much knowledge you had about each of the following topics before participating in this study. Press 1 to indicate that you had no knowledge about the topic and press 4 to indicate that you had a great deal of knowledge about the topic. If you had some knowledge about the topic press a number in-between.

APPENDIX G

EXPERIMENT 2 EVALUATION TASK

In this experiment you will be asked to evaluate the logical quality of several arguments. By 'logical quality' we mean whether the reason could provide support for the claim to a general audience. Thus, your job is to judge the extent to which the reason is structurally relevant to the claim--NOT whether or not you agree with the argument.

Before each trial you will see a screen that says "**NEW ARGUMENT**" On this screen be sure to position your hands on a comfortable home position on the keyboard with your thumbs on the spacebar.

The next screen will show you the claim of an argument. Read the claim as you would naturally and press the spacebar once you have finished reading. Then you will see a reason provided to support the claim. Read the reason as you would naturally and press the spacebar when you have finished reading.

Next you will be asked to rate the quality of the argument on a scale from 1(very low quality) to 8(very high quality). To rate the argument please use the number keys at the top of your keyboard.

REMEMBER: Your goal for this task is to judge the logical quality of the argument, not how much you believe it.

Evaluation Screen

Press a number 1-8 to rate the quality of the argument you just read

Very Low Quality Very High Quality

1 2 3 4 5 6 7 8

APPENDIX H

EXPERIMENT 2 CONTROL TASK

In this experiment you will be asked to read several arguments and then decide whether a word presented on the screen was a part of the argument you were just shown or not. While reading the arguments, please try to read them as naturally as possible. Do not try to guess the word we will ask you about while reading. Your task is to determine whether the word was present in just the most recent argument - NOT any other argument you read before.

During this task, please keep your hands in the home position on the keyboard with your pointer fingers on the F and J keys. You will be asked to press these keys to indicate whether you have seen the word shown. Press F to indicate that you HAVE seen the word and press J to indicate that you HAVE NOT seen the word presented.

REMEMBER: Your goal for the task is to judge whether or not you recognize the word shown from the argument.

Evaluation Screen

Press “F” if the word below was in the argument you just read and press “J” if it was not.

PROBE