The Influence of Controllability and Causal Status on Counterfactual Thinking and Associated Affect

by Mason R. Jenkins

B.A. in Psychology, University of North Carolina at Chapel Hill
B.A. in English, University of North Carolina at Chapel Hill
M.A. in Psychology, Northeastern University

A dissertation submitted to

The Faculty of
the College of Science of
Northeastern University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

June 11, 2019

Dissertation directed by

Nancy S. Kim
Associate Professor of Psychology
Dedication

I dedicate this dissertation to my grandmother Mavorine S. Jenkins for instilling a drive for a lifelong education.
Acknowledgments

I am forever indebted to many individuals who have supported my path in life. Without these people I would not be the researcher I am today.

I first thank Joseph Lowman for encouraging me to begin the very first inklings of a research career when he suggested I flesh out a proposal in his class into a full-scale research project. Many coffees were consumed in Davie Hall or on Franklin Street and they laid the foundation for the rest of my academic career. I thank Gunes Kutlu for getting me a foothold into post-baccalaureate research and Zach Rosenthal for allowing me to explore my emerging research interests in his lab.

Once in Dr. Rosenthal’s lab, I became fully aware that research is never an independent effort, and even if it could be, should not. I realized that no one person can ever possesses the perspectives required to understand all the different perspectives that can properly inform a project. I gained vast experience from Marissa Morris-Jones, Daniel Seed, and Melissa Sempowski about clinical concerns. I thank Jose Laurrai for his brevity and ability to articulate ideas about data, programming, and research in the comically succinct way that he can. I will forever be regretful when using a keyboard in Excel.

I owe a tremendous amount of gratitude to Lisalynn Kelley. She was my direct supervisor for over three years and the one primarily responsible for instilling in me a sense of research as a process, a method, a manner that needs to be followed. Very rarely is the scientific method most evident then when the data is in the process of being collected. In action, Lisa was always able to connect the dots, suss out what did and did not make sense, think two steps ahead, and identify unforeseen problems. However, no matter how you can boil a study protocol down to a decision tree, choices are never that simple and they require expert knowledge. Watching
her in action was to identify what kind of information you should know and what kind of researcher one should aspire to be. I only hope I represent Lisalynn well today.

When I came to Northeastern University, I did feel like I jumped from one moving train to another. My amazing advisor Nancy S. Kim helped to orient me to the new body of cognition literature and taught me to think as a researcher in a new domain. Together, we published studies and worked through many problems and issues in our research. Her help and critical eye on all matters was critical to the speed of my progress as a graduate researcher. With writing and presenting, she’s a tireless advocate of practice makes perfect and its truth. I’ve been an observer of this during my time in graduate school and I can only hope I continue to refine it as I continue in my career.

I do want to think two research sisters of mine, Caitlin Fang and Erienne Weine. I met Caitlin in my post-baccalaureate research career and followed her trajectory as a researcher as well. I had my first academic conference experiences with her and continue to cheer on her success. It has been valuable to have a colleague a few steps ahead of you in your career. I also want to think Erienne Weine for being a peer, friend, and cornerstone of my academic career.

Finally, the individuals outside the lab have also been critical to my success. I thank Zane Nichols and Lesley Seibel for helping establish a welcome home for me in Boston for many years. I also thank my partner Carlos Taylor for unequivocal support, love, and dedication throughout every step of this process. I also thank my parents, Wendy and Tracy, for the support they have given me along the way. I only hope with all of the above that I can repay in time what has been given to me.
Abstract of Dissertation

Counterfactual thinking involves mentally “undoing” an event and imagining how the resulting outcome could have been different. One body of existing literature provides a functional account of counterfactual thinking such that generating counterfactuals that are upward in nature (i.e., would lead to a more positive outcome) is thought to be largely adaptive in that it aids in planning and regulating future behavior, though at the cost of experiencing temporarily heightened negative affect (Epstude & Roese, 2008; Kahneman & Miller, 1986). It has been well established that the degree to which a given event is under the actor’s control (i.e., its controllability) predicts whether people mentally undo that event in counterfactual reasoning. With the notable exception of those suffering from severe depression, people tend to mentally undo past events that were within the actor’s control (Frosch, Egan, & Hancock, 2015; Markman & Miller, 2006), a finding consistent with the notion that upward counterfactual thinking may aid people in better planning their behavior in the future.

However, there are several major limitations to past research arguing for a link between counterfactual thinking and subsequent successful behavior or goal achievement. First, although counterfactual thought has a preparatory function for one’s own behaviors, past studies that experimentally manipulated the controllability of events to determine their subsequent influence on counterfactual thinking utilized only third-person hypothetical events (i.e., thinking about hypothetical others). Past results about third-person counterfactual thinking cannot automatically be assumed to be indicative of first-person counterfactual thinking given observers’ well-known attributional biases for others’ behaviors (e.g., the correspondence bias, Gilbert & Malone, 1995). Specifically, events perceived as highly controllable to external actors may be perceived as less controllable when imagining oneself as the focal actor. Because events
that are controllable are the likely focus of counterfactual thoughts, the results of such counterfactual thinking studies may differ radically in the first person.

Second, it has also been established that the location of a given event within the causal sequence of events that generated the outcome (i.e., its causal status; Segura et al., 2002) strongly influences whether people mentally undo that event. Yet whether controllability and causal status have interactive effects on counterfactual thinking remains unresolved.

Third, in past research it has been presumed that the controllability of events is the primary factor influencing the heightened negative affect accompanying upward counterfactual thinking (Mandel, 2003b; Markman & Miller, 2006). Yet given that causal status influences counterfactual thinking, it is reasonable to suspect that it may also influence the affective response, and causal status and controllability may also have interactive effects. Accordingly, the central goal of the current work was to address the following questions: (1) Do past findings on counterfactual thinking hold true when reasoning about oneself? (2) What are the separate and interactive influences of controllability and causal status on counterfactual thinking? (3) Does causal status (and controllability) predict changes in affect when people reason counterfactually about negative events in their own lives?

Experiments 1 and 2 found no difference between first- and third-person reasoners with respect to how controllability and causal status influenced their counterfactual thoughts. Experiment 3 established revisions to the methods used in past studies to cleanly test for an interaction of causal status and controllability. Experiments 4 – 6 provided evidence that people prefer to counterfactually intervene on intermediate rather than root causes in a causal chain, and suggest that causal status and controllability do not interactively influence counterfactual
thinking. These studies also suggested that the controllability of events may not, in fact, influence counterfactual thinking under the conditions tested.

Finally, Experiment 7 explored whether controllability and causal status predict changes in affect following upward counterfactual thinking about real first-person occurrences reported by study participants. The results of Experiment 7 showed, unexpectedly, that the influences of controllability and causal status on affect were moderated by domain. Causally deeper counterfactual thoughts predicted negative change in affect for interpersonal but not academic events. Generating counterfactuals for controllable events predicted positive change in affect for academic but not interpersonal events.

Together, these studies advance new knowledge about counterfactual thought and suggest that prior findings may not be as robust as previously assumed. Implications for these findings and suggested future directions in counterfactual and causal reasoning research are discussed.
Table of Contents

Dedication 3
Acknowledgments 4
Abstract of Dissertation 6
Table of Contents 9
List of Figures 10
List of Tables 11
Introduction 12
Experiment 1 31
Experiment 2 37
Experiment 3a 46
Experiment 3b 57
Experiment 4 63
Experiment 5 68
Experiment 6 74
Experiment 7 79
General Discussion 91
References 109
Footnotes 116
Tables and Figures 119
Appendix A 138
Appendix B 139
Appendix C 141
List of Tables

Table 1  128
Table 2  129
Table 3  130
Table 4  131
Table 5  132
Table 6  133
Table 7  134
Table 8  135
Table 9  136
Table 10  137
Introduction

Imagine arriving at the airport only to learn that you missed your flight by a matter of minutes. Almost immediately, you may think, “if only…” and complete that thought by mentally “undoing” a previous event that led to this outcome (e.g., if only I had taken a shortcut while driving, I would have made the flight; if only I had woken up earlier; if only I had booked a later flight; etc.). These “if only” thoughts are examples of counterfactual thinking, or the imagining of alternatives to reality by mentally “undoing” past events and imagining how the resulting outcome would be different.

The earliest research on counterfactual thinking referred to the process as a “simulation heuristic” that can facilitate inferences about the probable cause of an event (Kahneman & Tversky, 1982). For instance, given the above example, if we are able to mentally “undo” the act of waking up late and then simulate a subsequent sequence of events in which we arrive at the airport on time and do not miss our flight, then we are more likely to endorse waking up late as a candidate cause of missing the flight. Yet counterfactual thinking is not synonymous with causal inference. Consider the case of Mr. Jones (Kahneman & Tversky, 1982) who normally takes a direct route home from work every day. One afternoon was particularly pleasant and Mr. Jones chose to drive along a scenic route home in order to view the ocean. While going through an intersection along this route, a drunk driver charges through a red light and slams into Mr. Jones’s car, killing him. People who were asked to assume the perspective of Mr. Jones’s friends and family and complete the thought “if only…” were likely to counterfactually think, “if only he had taken his usual route home, he would still be alive”. However, when these same people were asked to identify the cause of Mr. Jones’s death, they identified the drunk driver as the cause rather than Mr. Jones’s decision to take the scenic route home.
To explain findings such as these, Mandel’s (2003a) judgment dissociation theory proposed that counterfactual thinking and causal inference are dissociable in that they ultimately differ in function. Whereas the goal of causal inference is to uncover the mechanistic explanation underlying an event, counterfactual thinking is concerned with identifying what could prevent an outcome from occurring again, regardless of whether it is part of a causal mechanism. In one experiment, participants read about Mr. Wallace, an influential member of an organized crime group who was poisoned and had his car forced off a cliff. Participants were likely to endorse the poison and car accident as causes of his death while thinking counterfactually that not being part of organized crime would be the best way to prevent his early death, rather than avoiding the poison or car accident. In particular, what they chose to counterfactually undo in this preventative manner tended to be an event they perceived to be within the actor’s control. For example, Mr. Wallace’s joining a life of organized crime may be perceived as relatively more within his own personal control than being poisoned by a vengeful enemy. In sum, people tended to think counterfactually about events that could prevent the outcome from happening again and that were relatively controllable by the actor.

A Functional Perspective on Counterfactual Thinking

Findings such as these are broadly compatible with the notion that counterfactual thinking can help people plan and regulate behavior for future events. If we think to ourselves while waiting in line at the airline ticket counter to reschedule our missed flight, “if only I had taken a shortcut while driving…”, we may be more likely to take such actions under similar circumstances in the future. Engaging in such counterfactual thought has been demonstrated to improve performance outcomes during laboratory tasks. Roese (1994) had participants complete two rounds of a difficult anagram task, with some participants engaging in a counterfactual
thought between the two rounds of the anagram task on how their strategies could be different and more successful. These participants ended up listing thoughts such as “if I looked for prefixes, suffixes and root words, I could get more correct answers”. The participants who did the counterfactual thinking task had higher accuracy scores in the second round of anagram puzzles compared to control counterparts who completed a filler task instead of the counterfactual exercise in between rounds. Similar results in task improvement were also found in people who were asked to land planes using a computerized flight simulator (Morris & Moore, 2000). Other studies have found that the effects of this counterfactual thinking intervention generalize outside the laboratory setting. In one study, college students who were asked to create counterfactual thoughts about how they could have done better on a recent test had a larger increase in grades later in the semester compared to students who were not asked to think counterfactually about their grades (Nasco & Marsh, 1999).

These findings have been collectively organized under a broad framework taking a functional perspective on counterfactual thinking (Epstude & Roese, 2008; Markman, Karadogan, Lindberg, & Zell, 2008; Roese, 1994; Roese & Epstude, 2017). Past research has shown that people have a tendency to think counterfactually about events with high controllability; that is, events that were within the actor’s control (Frosch et al., 2015; Markman & Miller, 2006). When people generate counterfactuals about more controllable than uncontrollable aspects of an event, they experience improvement on subsequent related tasks (Nasco & Marsh, 1999). Thus, counterfactual thinking may serve as a means by which future behavior can be more effective.

Despite the positive regulatory and preparatory influence that counterfactual thinking may have on behaviors, research has shown that these benefits may come at the price of
temporarily heightened negative affect. This appears to hold true when people generate so-called upward counterfactuals about how a situation could have turned out better (rather than when people generate downward counterfactuals about how a situation could have turned out more badly than it did).

Even the earliest research on counterfactual thinking appeared to presume that there must be a link between upward counterfactual thinking and experienced negative affect. Kahneman and Tversky (1982) initially described how the ease with which a counterfactual is judged to “undo” a negative outcome can lead to variations in the affective response to the event. They asked people to consider the following scenario:

“Mr. Crane and Mr. Tees were scheduled to leave the airport on different flights, at the same time. They traveled from town in the same limousine, were caught in a traffic jam, and arrived at the airport 30 minutes after the scheduled departure time of their flights. Mr. Crane is told that his flight left on time. Mr. Tees is told that his flight was delayed, and only left 5 minutes ago. Who is more upset?”

Kahneman and Tversky (1982) found near-consensus agreement that Mr. Tees would be more upset, with 96% of their participants choosing this traveler. Yet for both travelers the outcome was the same. Kahneman and Tversky (1982) argued that people chose Mr. Tees because it would be much easier for Mr. Tees to imagine a different upward counterfactual scenario that would have gotten him to the plane five minutes earlier than for Mr. Crane, who would need to have shaved thirty minutes off his travel time to have made the flight.

Thinking counterfactually about a serious negative event in one’s own life can induce more negative affect than attempting to explain how the outcome occurred. Mandel and Dhami (2005) asked prisoners to engage in either counterfactual or explanatory reflection over their own
actions leading up to and including their conviction and sentencing. Prisoners who were asked to think counterfactually about events (e.g., “If only I fled town after my crime, I would not have been caught”) experienced more guilt and self-blame for their actions than those who thought about their own actions from a non-counterfactual, explanatory perspective (e.g., “I was caught because I stayed too close to the scene of the crime”) with self-blame accounting for the influence of reflection type (i.e., either counterfactual or explanatory) on guilt.

Creation of upward counterfactuals about one’s own experiences has been linked to experienced regret as well as, less commonly, severe depressive tendencies. A recent meta-analysis offered a small but significant effect size for the association of upward counterfactual thinking and regret with depression symptoms (Broomhall, Phillips, Hine, & Loi, 2017). One study addressed whether upward counterfactual thinking would be disproportionately detrimental to the affect experienced by individuals with severe depressive symptoms compared to non-depressed and mild-to-moderate levels of depression. Markman and Miller (2006) had undergraduate participants describe a recent negative academic event and then create upward counterfactuals about how the event could have turned out better. They found that while all participants experienced an improvement in affect reflecting on the event after completing the counterfactual task, there was a significant difference in affect scores after the counterfactual task between participants who created a high compared to low ratio of controllable to uncontrollable thoughts. From this experiment, Markman and Miller (2006) concluded that focusing on controllable rather than uncontrollable aspects of a negative event may be a possible mechanism for improving affect after counterfactual thinking.

The idea that the relative controllability of events influences the affect experienced after counterfactual thinking was also supported in prior work. To experimentally test this, Mandel
(2003b) asked undergraduates to write about a recent negative event and then generate upward and downward counterfactuals about the event. Participants who were asked to write specifically about a recent negative academic event (an outcome which would be assumed to be more under the control of the participant, e.g., having not studied enough) experienced more negative emotion after upward counterfactual thinking than participants who wrote about an interpersonal event (an outcome which would be less fully under the control of the participant, e.g., the other people involved may be assumed to have also played key roles). To summarize the above findings, there is substantial evidence to support the notion that counterfactual thinking may have an adaptive preparatory function and is associated with heightened negative affect.

The Current Research: Understanding the Influence of Controllability and Causal Status on Counterfactual Thinking and Associated Affect

However, there are several important limitations to the past research I have outlined thus far. First, although a number of researchers have argued that upward counterfactual thought serves to guide one’s own future behavior toward achieving more optimal outcomes, past work on the influence of controllability on counterfactual thinking has utilized only third-person hypothetical events (i.e., thinking about hypothetical others) in experimental manipulations of event controllability. However, it may not be the case that these findings in third-person counterfactual thinking can be assumed to also appear in first-person counterfactual thinking. A wealth of literature on the correspondence bias indicates that when evaluating other actors’ behavior, observers tend to overestimate dispositional influences and underestimate situational influences on the actors’ behavior (Gilbert & Malone, 1995).

Second, it has also been established that the location of a given event within the causal sequence of events that generated the outcome (i.e., its causal status) strongly influences whether
people generate counterfactuals about that event (Segura, Fernandez-Berrocal, & Byrne, 2002; Wells, Taylor, & Turtle, 1987). Yet whether controllability and causal status have interactive effects on counterfactual thinking remains unresolved, as they have never previously been examined together. Third, in past research it has been presumed that the controllability of events is the primary factor influencing the heightened negative affect accompanying upward counterfactual thinking (Mandel, 2003b; Markman & Miller, 2006). Yet given that causal status influences counterfactual thinking, it is reasonable to suspect a relationship between the causal depth of generated counterfactuals and any change in one’s subsequent affective response.

Accordingly, the central goal of this dissertation is to address the following questions: (1) Do these past findings on counterfactual thinking hold true when reasoning about oneself? (2) What are the separate and interactive influences of controllability and causal status on counterfactual thinking? (3) Is there a relationship between the causal status of counterfactual thoughts (in addition to their controllability) and the changes in affect that are experienced after generating counterfactual thoughts about negative events in one’s own life? Below, I explain and justify these proposed research directions more fully.

**Reasoning about first-person versus third-person events.** Many experiments have demonstrated that observers tend to underestimate the role of situational factors and overestimate the role of dispositional forces when judging another person’s behaviors (Gilbert & Jones, 1986; Jones & Harris, 1967). This phenomenon is referred to as the correspondence bias (for a review see Gilbert & Malone, 1995). An initial demonstration of the correspondence bias was offered by Jones and Harris (1967). In their study, participants were shown an essay that either supported or opposed the then-president of Cuba, Fidel Castro. Some participants were informed that the writer picked what side to defend, thus making the essay a likely reflection of the
writer’s disposition and attitudes. Other participants were told that a debate coach determined the stance the writer took, making it impossible to know whether the essay reflected the writer’s personal views. Jones and Harris (1967) found that regardless of what the participants were told, the participants endorsed the view that the stance of the essay corresponded to the writer’s personal beliefs on Castro. That is, even participants aware of the debate coach’s influence still tended to think that the sentiments expressed in the essay reflected the writers’ own views and failed to adequately attribute those sentiments to the instructions of the debate coach. Jones and Harris (1967) concluded that observers fail to discount even highly salient situational factors when judging an actor’s behavior.

As previously discussed, many studies have examined the influence of controllable events on counterfactual thinking. Typically, they did so by experimentally manipulating controllability in third-person vignettes that participants read and for which they then generated counterfactuals. Given the seminal findings of Jones and Harris (1967) and subsequent replications (for reviews, see Gawronski, 2004; Gilbert & Malone, 1995; Malle, 1999), it is reasonable to hypothesize that the same events an observer judges to be influenced by the internal disposition of the focal actor (and thus relatively more controllable by the actor) may be judged to be influenced by the external situation (and thus relatively uncontrollable) when the observer instead imagines himself or herself as the focal actor.

To illustrate, recall Mr. Jones’s unusual drive home from work. Observers of Mr. Jones’s actions may attribute Mr. Jones’s unusual route along the shore to Mr. Jones’s disposition to go out of his way to enjoy scenery or perhaps to procrastinate doing chores at home. Thus, Mr. Jones’s decision is judged to be relatively controllable. Yet, it may be that if the observers themselves behaved similarly and took an atypical scenic route home, they would more often
attribute this to relatively uncontrollable situational forces (e.g., a hectic schedule or stresses in personal life that led them to take in a scenic view to relax).

Research consistent with a functional perspective on counterfactual thinking showed that upward counterfactuals tend to focus on controllable events, ostensibly better enabling one to regulate one’s own future behavior. It is critical to examine whether past findings are essentially replicated when reasoners adopt a first-person point of view. Take, for example, the studies of Frosch et al. (2015), which employed third-person point of view scenarios to determine if an event’s controllability interacts with its role as either an enabling condition or a cause of a negative outcome to influence counterfactual thinking. When the enabler and cause were mismatched in controllability, people generated counterfactuals about whichever event was controllable, regardless of whether it was an enabler or a cause. For instance, some people read about a student who could not be bothered to back up his thesis to the university server (i.e., a controllable enabler). The student then found that all of his computer files were corrupted when his virus checker malfunctioned and failed to incapacitate a virus embedded in an innocuous-looking email (i.e., an uncontrollable cause of the negative outcome). People who read this scenario tended to think counterfactually about the student’s controllable failure to back up files, even though other people had judged this event to be an enabling condition rather than a cause of the loss of the thesis (Frosch et al., 2015). Again, these findings are consistent with a functional perspective, but people in the study were making judgments about a third-person scenario.

I speculated that someone considering a similar situation in their own life may think quite differently about the situation. They may, for example, think, “the failure to back up my file itself did not cause the computer damage; the virus did the damage!” As such, their resulting counterfactual thought may instead be “if only the virus checker did not malfunction, I would not
have lost my thesis”. This situation-focused interpretation of events would also help preserve one’s view of the self, as it attempts to deflect any blame for the negative outcome. Prior studies have supported the notion that people tend to attribute positive outcomes to their own actions while attributing negative outcomes to external factors, a phenomenon known as the self-serving bias (De Michele, Gansneder, & Solomon, 1998; Miller & Ross, 1975; Shepperd, Malone, & Sweeney, 2008).

Theories of motivated reasoning may also predict differences in first- compared to third-person counterfactual reasoning for negative events. The Motivated Model of Hindsight Bias (Pezzo & Pezzo, 2007) contends that people engage in different reasoning strategies to avoid responsibility in the aftermath of negative outcomes (in ways they presumably would not be motivated to reason about external actors). Two ways in which people avoid responsibility are to view the negative outcome as unforeseeable or to view it as inevitable. As will be discussed in the introductions to Experiments 1 and 2, these two ways to avoid responsibility following negative outcomes might be expected to result in differences in how first- versus third-person reasoners counterfactually intervene on events. In sum, the literature on both the correspondence bias and motivated reasoning following failure suggest plausible hypotheses about how counterfactual thinking may differ between first- and third-person reasoners.

**Causal status.** Previous research showing that controllability influences counterfactual thinking has not yet examined whether the effects of controllability interact with any other major factors also shown to have an impact on counterfactual thought. Of particular interest in the current work is that the causal structure of an event can influence how people generate counterfactual thoughts about it. Segura et al. (2002) had participants engage in counterfactual thinking about scenarios described as a causal chain structure (see Figure 1) or as a conjunctive
common effect structure (see Figure 2). For example, their causal chain scenario described a shopper encountering a causally connected series of obstacles and setbacks on his drive to purchase a stereo system on the last day of a sale: a flat tire, a speeding ticket, a traffic jam, and a group of senior citizens crossing the road. The end result of this causal chain of events was that the shopper arrived to the store at a late hour, and the last stereo had been sold.

Participants were then asked to generate a statement about how the outcome could have been different. When the structure of the event was described as a causal chain (Figure 1), people were more likely to generate counterfactuals that undid the first event in the causal chain than any of the later events. Given that people undid the first event in the causal chain regardless of which event was described as occurring in that position, Segura et al. (2002) argued that causal status influenced counterfactual thought. Compatible findings were also found in earlier work, in which people judged events occurring earlier in the causal chain to be more mutable and thus presumably more ready for people to counterfactually undo (Wells, Taylor, & Turtle, 1987).

Segura et al. (2002) compared the above findings with people’s counterfactual thinking when presented with a second major type of causal structure. In their conjunctive common effect causal structure (Figure 2), the combined effect of four events collectively resulted in the effect, with no causal relationship between any of the four antecedent events. Four contestants were described as having been on a game show in which they could win a prize if three of them picked cards of the same color from their own individual standard deck of playing cards. Participants reading about this scenario learned that contestants each drew one card from their own deck in sequence. The end result was that two contestants picked black cards and two contestants picked red cards and thus did not win. Study participants, when asked to generate a counterfactual statement, were most likely to undo the action of the fourth and final contestant to alter the
outcome. Segura et al. (2002) concluded that the causal structure of a scenario predicts people’s counterfactual thinking about that scenario: People tended to mentally undo the initial event in a causal chain structure, and on the final event in a conjunctive common effect structure. These findings were not dependent on the number of events that precipitated the negative outcome, as people still chose to mentally undo the initial event when reading about a two-event causal chain and the final event in a two-participant card game.

The above influence of causal structure on counterfactual thinking may result from people’s overarching naïve theories about how outcomes occur. Segura et al. (2002) suggested that people mentally undid the first event in a causal chain because the resulting downstream events seemed less likely to occur without the presence of the first event. For example, they may have assumed that a prospective shopper would not have been likely to speed, thus getting the speeding ticket and being delayed long enough to then experience rush hour traffic, if he had not had a flat tire in the first place. In contrast, when choosing what event to undo in a conjunctive common effect scenario, people may have tended to treat the first three events as relatively fixed at the time when the fourth event occurs. Accordingly, they may have judged the final event (e.g., the fourth contestant’s card choice) to be the most changeable. However, people’s tendency to do so was attenuated by highlighting the mutability of the earlier events. Byrne, Segura, Culhane, Rasso, and Berrocal (2000) found that when giving people the same card picking game, and describing a technical hitch that forced the first contestant to redraw cards – thus picking a new, different colored card – participants become more likely to focus on this contestant in their counterfactual thinking compared to when no technical hitch was mentioned. Byrne et al. (2000) argued that the technical hitch story made the mutability of the first choice
more salient to participants, thus increasing the likelihood that they would generate counterfactuals about it.

However, although a number of studies have documented that causal status influences counterfactual thinking, exactly how it does so may depend on the domain. For example, people may not always reason counterfactually about initial (i.e., root cause) rather than intermediate events in a causal chain. Lagnado and Channon (2008) had people reason about everyday negative events that were organized in a five-event causal chain that always ended in a negative outcome. After reading the scenarios, participants made judgments of cause and blame for the second and fourth events in the causal chain. They found that there was a main effect of causal status for judgments of cause and blame, with the intermediate events (i.e., the fourth event in the chain) receiving higher ratings of causing the outcome and being to blame for the outcome compared to the more root events (i.e., the second event in the chain). Although Lagnado and Channon (2008) focused on judgments of cause and blame, their work provides reason to suspect that people’s counterfactual thoughts may not always focus on the root rather than intermediate causes in a causal chain. The reason for this is that people’s counterfactual judgments about these scenarios sometimes align with their causal judgments (Mandel, 2003a). When this occurs, we can expect people’s counterfactuals to “undo” more intermediate rather than root events. As Lagnado and Channon (2008) asked study participants to reason about a wide variety of everyday events, one possibility is that the preference for root events in the causal chain of Segura et al. (2002) and Wells et al. (1987) may be limited to the specific contents of the vignette they used (i.e., a shopper delayed on his way to the store), and may not reflect a more global counterfactual preference for root events.
An alternative possibility is that even when people are asked to generate counterfactuals about a wide variety of causal chains, that people will continue to prefer to “undo” root rather than intermediate events. Even though Lagnado and Channon (2008) found that intermediate causes received higher ratings of *cause and blame* compared to root causes, people’s *counterfactual thought* may not map directly onto these judgments. A divide between causal and counterfactual judgments would not necessarily be unexpected, as Mandel (2003a) argued that these judgments may not draw upon the same reasoning processes.

Although the above studies do not clearly establish where in a causal chain structure people generally choose to intervene, such studies have consistently shown that counterfactual thinking is influenced in some way by the causal structure of the scenario. That is, all the aforementioned studies show that whether people choose to mentally undo an event in that scenario depends on the causal status of that event in the causal structure.

**Causal status and controllability.** Do causal status and controllability interactively influence counterfactual thinking? Some clues may be derived from the studies by Frosch et al. (2015) described in the previous section, which examined whether the controllability of events interacts with causal status for simple causal scenarios (though these studies did not examine the more complex causal structures such as those found in Segura et al., 2002). Frosch et al. (2015) created 16 simple scenarios where a negative outcome was precipitated by two events. The first event was always a non-causal background enabling event that allowed the negative outcome to be able to occur. The second event was always a direct cause of the negative outcome. In both within- (Experiment 1) and between-subjects (Experiment 2) designs, Frosch et al. (2015) manipulated whether the first event and second event were controllable or uncontrollable (as pre-verified in a separate manipulation check). For example, one of their scenarios described a
student who had not bothered to take the time to back up his thesis to the university server. The virus checker on his computer became faulty and as a result his computer was attacked by a virus inside an otherwise normal-looking file. As a result, the student lost all his files including his thesis. This event described a negative outcome (computer files corrupted) that came after a first event (the enabling event: the thesis was not backed up) and a second event (a direct cause: getting the computer virus). In this scenario, not backing up the thesis enabled the negative outcome (yet did not cause it), as there was no direct causal link between not backing up the thesis and having your computer files corrupted. This event also contained a controllable element (not backing up thesis) and an uncontrollable element (virus checker became faulty and allowed a virus through). Other participants read a comparable scenario with the controllability of the first and second event reversed. In this scenario, the student had not been given access to the university server yet (an uncontrollable first event) and the computer files were lost after the student opened a suspicious email attachment from an unknown sender containing a virus (a controllable second event). Still other participants read about both a controllable first and second event or both an uncontrollable first and second event in this scenario.

Across two experiments, Frosch et al. (2015) had participants read the scenarios and complete the sentence, “Things would have turned out differently, if only…” as the key counterfactual generation task. The authors found that in these simple cause-effect scenarios, causal status interacted with controllability to influence the counterfactual thoughts generated. When the first and second event were mismatched in terms of controllability, participants’ counterfactuals focused on the controllable aspect of the scenario, regardless of whether it was the first or second event. However, when the first and second event were equally controllable or uncontrollable, participants’ counterfactuals were more likely to focus on the initial, enabling
event. A question addressed in the current work is how causal status and controllability may interact in more complex scenarios compared to the simple cause-effect scenarios tested by Frosch et al. (2015).

One additional reason to suspect an interaction of causal status and controllability in counterfactual thinking comes from Hilton, McClure, and Sutton (2010). They asked people to reason about causal chains containing events that were either voluntary human actions (e.g., tossing a lit cigarette into a bush) or natural, physical events (e.g., strong winds fanning the flames of a small fire) that precipitated a negative outcome (e.g., a forest fire). These intentional and natural events could also occupy either a root or intermediate position in the overall causal chain. Hilton et al. found that people attributed greater causal responsibility to intentional rather than natural causes and that this preference was qualified by an interaction with causal status. Specifically, intentional events were given more causal responsibility for the negative outcome when they were root rather than intermediate causes. Although intentionality is not the same construct as controllability, in this study the two constructs were confounded (i.e., voluntary human actions are more controllable than natural physical events). It is possible that an interaction of controllability and causal status could account for such findings. Additionally, although attributions of causal responsibility are not identical to counterfactual interventions, they may share a similar reasoning process (Lagnado & Channon, 2008). This evidence provides some reason to suspect a possible interaction of controllability and causal status in counterfactual reasoning.

**Controllability, causal status, and affect.** As we have just seen, the causal structure of an event has a significant influence on how people direct the focus of their counterfactual thinking (Lagnado & Channon, 2008; Segura et al., 2002). Yet it is currently unknown how
causal structure may influence the changes in affect associated with counterfactual thinking about an event. It may be the case that changes in affect differ depending on where people focus their counterfactual thinking in the causal structure of the scenario. That is, perhaps people experience different changes in negative affect when asked to think counterfactually about the initial cause versus the last cause in a causal chain sequence precipitating an outcome.

To illustrate, consider again the earlier example of the shopper who encountered a causal chain of impediments on the way to purchase a stereo on sale. In this example, the flat tire was the first event that launched the causal chain of subsequent delays, ending with waiting for the group of senior citizens to cross the street. Perhaps if this shopper generated a counterfactual about the first event (in this case, the flat tire), they would experience more negative affect than they would when focusing on the final event (in this case, the elderly group). Thinking counterfactually about the flat tire may make the shopper more acutely aware of all the events that transpired to thwart the shopper from purchasing the stereo, thus increasing the negative affect the shopper might feel. As such, thinking counterfactually about the final event may not result in as much negative affect, as only one event before the outcome has been made salient.

On the other hand, it may alternatively be the case that thinking counterfactually about the final cause leads to an increase in negative affect compared to thinking about the initial cause. In the same way that Mr. Tees feels worse learning he just missed his plane by five minutes compared to Mr. Crane who missed his plane by 30 minutes, perhaps thinking counterfactually about the last event preceding the negative outcome engenders more negative affect than the root cause as it is easier to imagine simply avoiding the group crossing the street compared to avoiding all the events resulting from the flat tire. Given these possibilities, causal
structure may therefore be a reasonable candidate factor influencing changes in affect after upward counterfactual thinking.

In complex events, it is also possible that causal status and controllability interactively influence the negative affect that people experience after upward counterfactual thinking. Recall the scenario involving the shopper who arrived too late to purchase a stereo on the final day of a sale. This event was characterized by a four-event causal chain of events precipitating the negative outcome. For this event structure, the earlier events had a higher causal status than events occurring later, closer to the final outcome. Suppose that the shopper thinks counterfactually about a controllable event with a high causal status. For instance, he could think counterfactually about his initial decision to take a detour, which caused him to drive over broken glass and get a flat tire. Generating a counterfactual about this controllable event with a higher causal status may lead to greater subsequent negative affect. Now, consider the case in which the first event in this same chain of events was instead uncontrollable. Suppose, for example, that the detour that the shopper took was the result of road work and was thus a necessary detour. Thinking counterfactually about this event may not engender as much of an increase in negative affect.

If we go on to consider the case of controllable events with lower causal status, thinking counterfactually about these may cause less negative affect compared to thinking counterfactually about the same events when said to be uncontrollable. As an example, imagine the final event the shopper experienced before the outcome was the delay from the shortcut taken. Even though this event is controllable, it may not lead to as much negative affect. Based on the setbacks occurring earlier, the shopper may not perceive this event as being able to solely undo the negative outcome. On the other hand, perhaps thinking counterfactually about an
uncontrollable final event preceding the negative outcome creates more negative affect. In this case, the shopper becomes frustrated thinking that despite all of his controlled actions, something unexpected and uncontrollable at the very end is what prevented him from getting his stereo. In sum, then, it seems plausible that the causal status and controllability could have an interactive effect on the temporarily heightened negative affect following upward counterfactual thinking.

**Overview of Experiments**

In Experiments 1-2, I addressed the influences of causal structure and controllability on the generation of counterfactual thoughts when reasoning about first-person hypothetical events. As past work on the influence of causal status and controllability on counterfactual thinking has been limited to thinking about third-person hypothetical events (i.e., reasoning about hypothetical others), the first two experiments tested whether past findings on counterfactual thinking extended to when people reasoned counterfactually about themselves. In Experiment 3, I identified some confounds in scenarios used in the seminal past work on counterfactual reasoning for causal chains, which would make it difficult to clearly test whether the causal status and controllability of events have an interactive effect on the formation of counterfactual thoughts. Given the limitations identified in seminal materials in Experiment 3, in Experiments 4 through 6, I used a variety of stimulus materials, experimental designs, and coding methodologies to thoroughly examine whether the causal status and controllability of events have separate or interactive effects on counterfactual thinking. Finally, Experiment 7 was an exploratory study that employed participants’ reports of their own personal events to test whether causal status and controllability influence counterfactual thoughts and subsequent changes in affect. Most events in daily life rarely have the simplistic causal structures employed in controlled laboratory studies. Whereas Experiments 1 through 6 involved thinking about causal
structures such as the ones displayed in Figures 1 and 2, people’s own mental models are unlikely to isolate the individual events comprising the scenario in such a clear-cut, discrete fashion. The causal structures of people’s real-life personal events may be considerably more complex than the earlier described causal structures, more akin to a “causal web” (Strickland, Silver, & Keil, 2017). Therefore, in Experiment 7, causal mapping techniques were employed to identify the causal status of events and counterfactuals.

Experiment 1

As mentioned earlier, past research testing how various factors influence counterfactual thinking has primarily employed third-person hypothetical scenarios rather than having people reason about how they themselves would think and reason given a negative event. Given the well-established finding that people may reason differently about themselves versus others (e.g., the correspondence bias), Experiments 1 and 2 were designed to test whether past work on the influences of event controllability and causal status (respectively) on counterfactual thinking hold true for both first- and third-person reasoners.

Experiment 1 addressed whether people are more likely to generate counterfactuals about more controllable than uncontrollable events, regardless of whether they are reasoning from the first- or third-person point of view. As described in the Introduction, the correspondence bias describes how the behavior of a third-person actor may be viewed as more intentional and less explained by situational factors compared to when people reason about the cause of an event occurring to themselves (which they would be more likely to attribute to situational factors). Thus, people may generally believe that events that happen to others are within that actor’s control, while the same events occurring to the observers themselves would be judged to be relatively less within their own control.
In addition to the correspondence bias, research on the topic of motivated reasoning provides additional evidence to suspect that patterns of counterfactual thinking may differ depending on point of view. The Motivated Model of Hindsight Bias (Pezzo & Pezzo, 2007) contends that people engage in different reasoning strategies to avoid responsibility in the aftermath of negative outcomes. One way that people avoid responsibility is to view the negative outcome as having been unforeseeable. In addition, when people view negative outcomes as inevitable (and are not motivated to think differently about other people’s negative outcomes), then event controllability may not influence counterfactual thought at all, as people may not view the events precipitating their failure as within their control. Thus, the literature on both the correspondence bias and motivated reasoning following failure suggest plausible hypotheses about how counterfactual thinking may differ, if at all, between first- and third-person reasoners.

Experiment 1 is a replication and extension of past work contending that event controllability influences counterfactual thought (Frosch et al., Experiment 2, 2015). Based on prior work, I hypothesized that people’s counterfactual thoughts would generally continue to focus on controllable rather than uncontrollable events. More importantly, by having people reason about events from either the first- or third-person perspective, I also sought to examine whether point of view moderates any influence of event controllability on counterfactual thinking.

Method

Participants. I recruited 148 adult participants on Amazon Mechanical Turk (\(M_{age} = 34.96, SD = 10.58; 47.5\% \text{ female} \)) to have a power of at least .80 based on the between-participant controllability effect size of \(d = 1.28 \) found in Frosch et al. (2015). Participants
received $1.00 for their participation, an amount previously shown to incentivize participants to complete a fifteen-minute task (Buhrmeister, Kwang, & Gosling, 2011).

**Materials.** I used the eight scenarios from Frosch et al. (2015; Experiment 2) that, in their pretesting, showed the highest agreement rates for judgments of controllability and identification of enablers and causes. Since I was primarily interested in the competing effects of controllable and uncontrollable antecedents to the outcome in this study, I did not use the additional scenarios from Frosch et al. (2015) in which the level of controllability of the cause and enabler were equated. As only U.S. participants were recruited, I changed several nouns from their original British English to their American English counterparts to prevent confusion (e.g., changed “lorry” to “truck”; “lift” to “elevator”).

I manipulated point of view by changing the pronouns in the scenarios between conditions. For the third-person condition, I kept the same pronouns and names as originally used in Frosch et al. (2015). In the first-person condition, names, pronouns, and verbs were changed in the scenarios to be written from a first-person past perspective. This allowed participants to imagine themselves as the subject who had experienced these events and the negative outcome. Any reference to an individual other than the first-person subject was changed from a proper name to a generic name (e.g., “Phoebe driving her van” to “a driver in a van”), because the names of strangers would only have been known in the third-person omniscient perspective adapted by Frosch et al. (2015), and would not have been known at the time from a first-person perspective.

**Design.** Point of view was manipulated between participants, such that each participant either read eight third-person scenarios or eight first-person scenarios. Event controllability was manipulated within participants. For all scenarios, the first event described was a background
event that allowed the negative outcome to occur (e.g., important files had not been backed up to a server) and the second event was a direct cause of the outcome (e.g., a virus corrupted all the files on the computer). For each scenario, the first and second events each had a controllable version and an uncontrollable version (e.g., in the uncontrollable version, computer files were not backed up because the student had not been given access to the university server; in the controllable version, the student could not be bothered to take the time to back up the files). The results of a stimulus norming task in Frosch et al. (2015) confirmed that the two types of events were judged to be either controllable by the focal actor or uncontrollable, as intended. Each participant saw four scenarios where the first event was controllable and the second event was uncontrollable, and four scenarios where the first event was uncontrollable and the second event was controllable. The specific scenarios were counterbalanced between participants such that the four scenarios viewed by one half of the participants as first event - uncontrollable, second event - controllable would be viewed as first event - controllable, second event - uncontrollable by the remaining participants. The order in which scenarios were presented was randomized for each participant.

**Procedure.** All participants completed the study online using the Qualtrics software survey tool (Qualtrics, Provo, UT). For each scenario, participants were instructed to read it while imagining themselves experiencing the scenario. After reading each scenario, participants were asked to complete the sentence, “Things could have turned out differently, if only…” and were given an open box to type their response. This constituted the counterfactual thinking task. One scenario and response box was presented on screen at a time.

After reading all eight scenarios, participants completed a reading comprehension question where they were asked to identify which event was not part of any of the prior
scenarios. This question served as a basic check on their attentiveness during the study. Participants then completed demographic questions and were informed about the purpose of the study.

**Results and Discussion**

The data from 13 participants (8.78%) was excluded from the main analyses because these participants incorrectly answered the reading comprehension check at the end. As in Frosch et al. (2015), two raters independently coded each counterfactual statement regarding whether it referred to the first event, the second event, both the first and second event, or other. The raters had an agreement rate of 90.00%, which matched that of the original study. Disagreements were resolved through discussion. For each participant, two scores were then calculated. One was the percentage of all scenarios where the counterfactual referred to only the first event, and the other was the percentage of all scenarios where the counterfactual referred to only the second event.

I employed both frequentist and Bayesian analytic approaches to help adjudicate whether there is evidence in support of differences as well as support of similarities between the factors of interest throughout the main experimental tests in the current work. Bayes factors (BFs) can represent the ratio of evidence for the alternative hypothesis over the null hypothesis (denoted as BF_{10}) and thus also represent the ratio of evidence for the null hypothesis of no differences between means (represented as H_{0} throughout the current work) over the alternative hypothesis of differences between means (H_{1}). For instance in Experiment 1, support for H_{0} would mean that the variable did not influence people’s counterfactual thought while support for H_{1} would mean that the variable did influence people’s counterfactual thought. Both frequentist and
Bayesian analyses are reported in the text and in tables following approaches used by other researchers identifying similarities and differences (e.g., Livingstone & Isaccowitz, 2019).

Frequentist analyses were conducted using SPSS 25 (IBM Corp., 2017) and Bayesian analyses were performed with JASP v. 0.9.2 (JASP Team, 2018). JASP yields a BF for the inclusion (BF_{incl}) of each effect in the overall model (i.e., support of a main effect or interaction) using a default set of priors. The interpretive guidelines for the BFs are those provided by Lee and Wagenmakers (2013). Bayes factors ranging from 3 to 10 represents “moderate” evidence in support of H$_1$ over H$_0$, a BF ranging from 10 to 30 represents “strong” evidence, a value from 30 to 100 is “very strong”, and a BF over 100 is “extreme”. For evidence supporting H$_0$, a BF between .100 and .333 is “moderate”, between .033 and .100 is “strong”, between .001 and .033 is “very strong”, and a value less .001 is “extreme”. Values of BF between .333 and 3 are inconclusive for either hypothesis.

I conducted a 2 (event number: first event, second event) x 2 (controllability: controllable, uncontrollable) x 2 (point of view: first-person, third-person) mixed-factors ANOVA on the percentage of scenarios receiving the focus of counterfactual thought for each of the first events and each of the second events, following the analysis procedure used in Frosch et al. (2015).

Frequentist and Bayesian analyses are summarized in Table 1. There was a main effect of event controllability, $F(1, 126) = 62.61, p < .001, \eta^2 = .33, \text{BF}_{\text{incl}} > 1000$. People generated counterfactuals about controllable events in a higher percentage of scenarios ($M = 27.53\%, SD = 12.45$) compared to uncontrollable events ($M = 13.34\%, SD = 14.14$). There was no main effect of event number, $F(1, 126) = 3.51, p = .064, \eta^2 = .03, \text{BF}_{\text{incl}} = .255$. There was no interaction of event number and controllability, $F(1, 126) = 1.87, p = .174, \eta^2 = .02, \text{BF}_{\text{incl}} = .425$. There was
no main effect of point of view; nor was there an interaction of point of view and any within-participant factor on the percentage of scenarios receiving counterfactual focus (all $Fs < 1.90$, $ps > .170$, $\eta_p^2 < .02$, $BF_{incl} = .055$).

In sum, Experiment 1 replicated past findings showing that people are generally more likely to generate counterfactuals referring to controllable rather than uncontrollable events. I also found that they preferred to generate counterfactuals about the initial, enabling event that was shown in the original study (Frosch et al., 2015). This finding was also compatible with prior work arguing for a counterfactual preference for initial events in a sequence (Segura et al., 2002; Wells et al., 1987). The current experiment, however, provided evidence against an interaction of perspective (in terms of first- versus third-person point of view) and event controllability on people’s counterfactual thinking, suggesting that point of view does not moderate the previously established effect of controllability on counterfactual thought.

**Experiment 2**

Experiment 1 produced no evidence that point of view moderates the previously established influence of event controllability on counterfactual thinking. Experiment 2 was designed to test whether past findings regarding the influence of causal structure on counterfactual thinking held true when people were asked to imagine the event happening to themselves as well as to others. Specifically, the causal structure and counterfactual thinking study of Segura et al. (2002) was modified in Experiment 2 to examine people’s counterfactual thinking in the hypothetical first (instead of third) person.

I hypothesized that evidence for a self-serving bias may lead first- compared to third-person reasoners to counterfactually intervene on different nodes in the causal structure of an event. Drawing from the Motivated Model of Hindsight Bias (Pezzo & Pezzo, 2007) discussed
earlier, one way that people avoid responsibility for failure following a negative outcome is to view the negative outcome as unforeseeable, even in hindsight. It is possible that when people are motivated to view their negative outcomes (e.g., experiencing disappointment and failure) as unforeseeable, they may tend to counterfactually intervene on relatively more causally intermediate events than root-cause events in a causal chain of events (Figure 1 depicts an example of a causal chain structure). That is, when people are motivated to reason that there is no way that they could have anticipated a negative outcome given the occurrence of the causally deeper events, people’s counterfactuals might be less likely to focus on those causally deeper events. It is possible that such a phenomenon may also influence reasoning about events in a conjunctive common-effect structure. Segura et al. (2002) found that third-person reasoners generated more counterfactuals about the final event than the first event in a conjunctive common-effect structure (Figure 2 depicts an example of a conjunctive common-effect structure). The motivation to view the outcome as unforeseeable may attenuate any tendency to judge the final event (or any singular event) as more influential on bringing about the outcome compared to the prior events. Thus, to the degree that people’s motivations to infer foreseeability differ between the first and third person, point of view may moderate the extent that causal structure influences the location of counterfactual thought.

**Method**

**Participants.** Seventy-one Northeastern undergraduate students ($M_{\text{age}} = 18.82, SD = 1.06; 53.5\%$ female) received partial course credit for their participation. This sample was powered over $.80$ based on the causal status effect size ($d = .52$) identified in Segura et al. (2002).
**Materials.** I created two sequences of events to correspond to the causal chain (Figure 1) and conjunctive common effect (Figure 2) causal structures used in prior research on causal structure and counterfactual thinking (Segura et al., 2002; Wells et al., 1987). Each sequence included four events culminating in receiving a bad grade on a test (i.e., the outcome). A written description of the events was presented to participants along with a visual diagram depicting the same sequence of events for each of the two types of causal structures.

I manipulated between participants whether scenario contents were written from the first- or third-person perspective. In the first-person perspective, participants were instructed to read each sequence and to imagine themselves being in that situation. For the causal chain structure, the event given was as follows: “Before an upcoming test, you caught a cold and became sick. Being sick, in turn, caused you to get little sleep. Getting little sleep made you have problems managing your time and responsibilities. These difficulties in managing your time and responsibilities caused you to miss a review session for your coming test. The end result of all of these events is that you did poorly on the test.” Figure 3a displays the visual diagram reflecting this causal chain scenario.

For the conjunctive common effect structure, the event was described as follows: “Upon arriving to a class to take a test, the following events occurred. As you entered, you realized that the room was too hot, which made you uncomfortable. You then chose a desk and sat in it, but it had a wobbly seat, which was bothersome. Then, you noticed that there were loud noises from a construction site outside. Finally, you realized that you were much hungrier than you normally would be during class. The end result of all of these events is that you ended up doing very poorly on the test.” Figure 3b displays the visual diagram reflecting this conjunctive common-effect scenario.
For the third-person condition, the scenarios were identical to that of the first-person condition except that all pronouns were converted to the third-person perspective. Participants were asked to imagine the events described as happening to two hypothetical female college students. The names chosen (Lisa and Karen) were equated in terms of inferences people have been shown to draw from them with regard to their perceived age, attractiveness, and competence in prior work (Kasof, 1993).

For both sequences, participants were also told to assume that each event described would have been equally bothersome or distracting to them or to the focal individual. They were also instructed to assume it was the first time that any of the events described occurred to them or the focal actor while they had been enrolled in that course. These details were included to prevent participants from focusing their counterfactual thoughts on rare or uncommon events (cf. Kahneman & Miller, 1986; Kahneman & Tversky, 1982) or based on idiosyncratic notions of causality based on preference. Additionally, when reading the conjunctive common effect sequence, participants were told to assume that “one, two, or three of these events would not have been enough to cause you [Karen] to do poorly on the test. It was the presence of all four events together that caused you [Karen] to do poorly”. This was added to clarify the precise causal structure of the conjunctive common effect sequence.

Finally, to rule out any effects of content, an additional condition was created for each causal structure, in which the sequence of events was described in reverse order. For example, the reverse description for the causal chain event was “Before an upcoming test, you [Lisa] missed a review session for the test. This, in turn, caused you [Lisa] to have problems managing your [her] time and responsibilities, as you [she] had to make time to compensate for the missed review. These scheduling problems caused you [Lisa] to get less sleep. Getting less sleep
caused you [her] to catch a cold and become sick. The end result of all these events is that you [she] did poorly on the test.” The order of events in the conjunctive common effect sequence was reversed as well. The entire study was programmed using the online Qualtrics survey design suite (Qualtrics, Provo, UT).

**Design.** A mixed factors design was employed. Point of view was manipulated between subjects with 36 participants in the first-person condition and 35 participants in the third-person condition. Causal structure was manipulated within subjects; all participants viewed both a causal chain story and a conjunctive common effect story. The order of presentation of causal structure was counterbalanced between subjects. Event sequence was also manipulated between-subjects; participants were presented with events occurring in either the original or reversed order.

**Procedure.** After reading each sequence of events and viewing the visual diagram, participants were asked, “which of the events would you choose to ‘undo’ and cause to be different in order to have done better on the test?” Participants were then prompted to indicate their choice of event from among the events depicted in the causal diagram. After making their choice, participants were asked, “How plausible do you think it is that something like the prior event could happen?” on a 7-point scale (1 = not plausible at all; 7 = very plausible).

**Results**

Plausibility scores for the forward and reverse sequences were collapsed for each causal structure, and a paired samples t-test revealed no difference in plausibility scores between the causal chain structure ($M = 4.72; SD = 1.81$) and conjunctive common effect structure ($M = 4.36; SD = 1.77$), $t(35) = 0.97$, $p = .338$. 
A chi-square test of goodness of fit was conducted to examine the distribution in frequencies for people’s preferences for undoing each of the four possible events in each sequence as a function of point of view. Bayesian independent multinomial tests were conducted to provide the corresponding BF s. A 2 (condition: first, third-person) x 4 (node: 1, 2, 3, 4) chi-square test revealed no effect of point of view on which node in the causal chain people preferred to counterfactually “undo,” $X^2 (3, N = 71) = 3.97, p = .265, BF_{10} = .198$. Similarly, there was no effect of point of view on which node was chosen in the conjunctive common-effect scenario, $X^2 (3, N = 71) = 0.69, p = .876, BF_{10} = .089$.

Given that there was no effect of point of view on the counterfactual choices for each of the causal structures, I collapsed counterfactual responses across point of view to see if the prior findings on how causal structure affects what nodes people prefer to counterfactually “undo” still emerge in the current study. Table 2 displays the number of participants who chose to counterfactually “undo” each node by condition for each causal structure. A 2 (causal structure: causal chain, conjunctive common-effect) x 4 (node: 1, 2, 3, 4) chi-square test revealed that counterfactual node preference depended on which causal structure it was in, $X^2 (3, N = 71) = 19.51, p < .001, BF_{10} = 377.100$. To further understand this dependency, two separate chi-square goodness of fit tests and corresponding Bayesian binomial tests were conducted to see whether the distributions in counterfactual choices for the two causal structures differed from a chance distribution. The pattern was such that people most frequently preferred to “undo” the first (root) node in the causal chain structure ($BF_{10} > 1000$) compared to all the other subsequent intermediate causes $X^2 (3, N = 71) = 44.78, p < .001$. Distribution was above chance for the first node ($BF_{10} > 1000$) and below chance for the third and fourth nodes ($BF_{10} = 192.449$ each). There frequency for the second node was similar to chance, $BF_{10} = .260$. However, they showed
Discussion

The results of Experiment 2 showed that when people are asked to reason about a causal chain scenario resulting in a negative outcome, they prefer to counterfactually “undo” early events in the causal chain rather than later events. This preference occurs regardless of whether people reason about first- or third-person events. Thus, there is no evidence for a correspondence bias in reasoning counterfactually about causal chain scenarios.

Additionally, people showed no preference to undo any event in a conjunctive common effect scenario. This lack of preference to undo events in any particular position in this causal structure was found regardless of whether people were asked to reason in the first or third person. This pattern of results differed from prior work on third-person counterfactual thinking on conjunctive common effect scenarios (Segura et al., 2002). It is possible that the preference seen in prior work for people to counterfactually undo the last event in a conjunctive common effect was specific to the domain of the event used in that work (i.e., drawing from a deck of playing cards), and may not be extrapolated across domains (e.g., to academic performance). Regardless, the results suggest that when the domain of an event is held constant across the first and third-person perspectives, people’s counterfactual reasoning appears to remain unchanged.

Interim Summary of Experiments 1 and 2

Experiments 1 and 2 extended past work on counterfactual thinking by demonstrating, for the first time, that event controllability and causal status systematically influence counterfactual thinking about the self as well as about others, at least with the kinds of scenarios used in past work on reasoning in the third person\(^4\). Because Experiments 1 and 2 demonstrated that the

\[ \text{BF}_{10} < .198, \chi^2 (3, N = 71) = 1.28, p = .733. \]
influence of controllability and causal status on counterfactual reasoning did not differ depending on point of view, in the next set of experiments I was able to focus solely on counterfactual reasoning for first-person scenarios. In Experiments 3a–6, I asked participants to reason about hypothetical first-person scenarios to determine whether or not controllability and causal status interactively influence counterfactual thinking.

**Introduction to Experiments 3a–6**

The central purpose of Experiments 3a through 6 was to test for such an interaction. As I discussed in the introduction of this paper, prior research provides some reason to expect an interactive influence of causal status and event controllability on people’s counterfactual thinking. In Frosch et al. (2015), people were asked to generate counterfactuals about scenarios consisting of two events that precipitated a negative outcome: An enabling event and a directly causal event. A variety of scenarios were used, and versions of each scenario were written so that the enabling event and the cause could be construed (and were reliably judged) to be either controllable or uncontrollable. The authors found that when the enabler and cause varied in their level of controllability, people’s counterfactuals were more likely to focus on the more controllable event, regardless of whether it was an enabler or a cause. Yet, when the two events were similar in their level of controllability, there emerged a preference for the initial, enabling event. Although this study did not manipulate causal status of events in a causal structure per se, it did indicate that controllability and cause-related role (i.e., enabler versus cause) can interactively influence the focus of people’s counterfactual thought.

Additionally, Hilton et al. (2010) found support for an interaction of causal status and level of human involvement with an event (i.e., whether an event was an intentional human action or an act of nature or natural forces) in judgments of causal responsibility. When people
reasoned about causal chains of events, wherein each of these events occupied either a root or intermediate causal position, people’s responsibility judgments focused on the intentional human actions more than the natural events; furthermore, intentionality and causal position had an interactive effect. Intentional human actions were rated as more causally responsible in bringing about the negative outcome when they occurred as root rather than intermediate causes in the overall causal chain. Hilton et al. (2010) manipulated intentionality rather than controllability and measured causal judgments rather than counterfactual judgments. Yet their work suggests that forces of human agency (which are controllable relative to forces of nature) can interact with causal status to influence causal attributions. Given that counterfactual intervention sometimes aligns with causal judgment (Mandel, 2003a), it is plausible that similar findings might be obtained for counterfactual thought as well.

To systematically test whether controllability and causal status interactively influence counterfactual thinking, I conducted a series of interrelated experiments. First, Experiments 3a and 3b were aimed at developing and testing stimulus materials that were balanced for possible confounds not addressed in previous seminal work -- such as unequal causal strengths between links in the chain and imbalances in judged event controllability across causal positions – that could then be used to cleanly test for an interaction in subsequent experiments. In Experiment 4, I used the balanced set of stimulus materials developed in Experiment 3b in a counterfactual reasoning task to conduct the first direct test for an interaction of causal status and controllability. Experiment 5 was then designed to determine whether the results found in Experiment 4 would be replicated and whether they are robust across different types of counterfactual elicitation tasks common to the literature. As such, in Experiment 5, people were prompted with one of two common types of counterfactual elicitation methods from prior
research to see whether findings would be consistent across them. Experiment 6 was designed to have participants, rather than external raters, classify the focus of their own counterfactual responses. Together, the results of Experiments 4–6 were designed to yield information about whether there is an interaction of causal status and controllability on counterfactual thought, and whether or not the results are consistent across a variety of methodological approaches used in the counterfactual reasoning literature.

**Experiment 3a**

In Experiment 3a, I examined whether the legacy stimulus materials from Wells et al. (1987) and Segura et al. (2002) – wherein the events in a scenario are rotated between-subjects throughout all causal positions – were controlled for independent judgments of the causal strengths of events and of the controllability of those events. In addition, across Experiments 3a and 3b, my eventual goal was to assemble a set of well-balanced materials that could be used to test for an interaction of causal status and controllability in counterfactual reasoning.

Prior work on counterfactual thinking has generally employed one of two experimental approaches to construct and present hypothetical scenarios, as reflected in Experiments 1 and 2 in the current work. The method used by Wells et al. (1987) and Segura et al. (2002) was to generate a single base story with multiple causal links and events that could be manipulated between conditions. Recall that Segura et al. (2002) and Wells et al. (1987) asked people to reason counterfactually about a scenario in which a hypothetical shopper, William, was delayed by obstacles on his drive to purchase one of a limited number of stereos on sale at a store, with the end outcome that he arrived at the store after the last stereo had been sold. The four obstacles were receiving a speeding ticket, getting a flat tire, being caught in a traffic jam, and having to wait for a group of slow senior citizens to cross the road. The events were written to form a
causal cycle of events, such that each event could be rotated throughout all four causal positions in a causal chain (e.g., root cause, intermediate causes) to form different story versions presented between subjects without affecting the coherence of the story.\textsuperscript{5}

It was critical that I test for several potential confounds in these stimulus materials. First, the original scenario stated that William’s intentional decision to drive fast caused him to receive a speeding ticket, whereas the delays from the flat tire, the traffic jam, and the senior citizens were described as unexpected setbacks (Segura et al., 2002; Wells et al., 1987). Thus, in the original scenario, causal status and controllability may have been confounded; three of the four events were likely judged to be uncontrollable while only one event (the speeding ticket) was likely judged to be within William’s control. Given people’s preference to counterfactually undo controllable rather than uncontrollable events, it is unknown whether the preference to undo root rather than downstream causes in a chain found by Segura et al. (2002) might be driven or further qualified by some interaction of causal status and controllability.

A second concern regarding prior work is that the strengths of the causal relationships between the events are unknown. Although the events were described in a causal chain, previous work did not control for whether some links were judged to have a stronger cause-effect relationship than others. For instance, it is possible that people may judge the delay from changing the flat tire to strongly cause William’s having entered the start of the rush hour traffic jam, whereas people may not judge the delay from waiting for senior citizens to cross the road to have the same strength of causal impact on William’s intentional decision to speed. Controlling for the strengths of the causal links between events in the scenario is potentially important because people might prefer to, for example, counterfactually undo strong causes over weak causes. In sum, it is unknown whether people’s previously documented preference to undo root
rather than intermediate causes in a causal chain would continue to emerge when event controllability is controlled for and causal strengths of the links between events are balanced throughout all the causal positions in a scenario.

Experiments 3a and 3b were therefore aimed at providing information about the judged causal strength and controllability of (1) the original stimulus materials in Segura et al. (2002) and Wells et al. (1987), and (2) additional variations of those materials that, taken together, could ultimately enable me to test for an interaction of causal status and controllability (i.e., in Experiments 4-6). Given that the original scenario in Segura et al. (2002) and Wells et al. (1987) was comprised of one relatively controllable event (speeding ticket) and three relatively uncontrollable events (flat tire, traffic jam, seniors crossing the road), I constructed additional stimulus materials such that the speeding ticket could be potentially judged as uncontrollable (e.g., caused by a fallen tree limb blocking a sign signaling a drop in speed limit), and the three uncontrollable events could be construed as being within one’s control (e.g., intentionally taking a shortcut over a very rough gravel road and receiving a flat tire). I asked participants to read the individual events or causal pairs, imagine the events occurring to them, and then rate their beliefs about either the causal relationship between events or their degree of perceived controllability.

To recap, in Experiment 3a, I documented the extent to which the original stimulus materials in Segura et al. (2002) and Wells et al. (1987) were balanced in terms of event controllability and causal strengths of the links between events. In addition, in Experiment 3a, I measured the judged event controllability and causal strengths of the links between events for additional variations on the materials, with the ultimate goal of forming a full set of scenario materials balanced for causal strength and cleanly manipulated for controllability to use in Experiments 4-6. As Segura et al. (2002) found similar patterns of counterfactual reasoning for
two- as well as four-event causal chains, I also aimed to assemble balanced stimulus sets
containing non-overlapping contents (i.e., two-event scenarios with the speeding ticket and flat
tire events and two-event scenarios with the traffic jam and senior citizens).

Method

Participants. I recruited 49 Northeastern University undergraduates to take part in this
task for partial course credit ($M_{age} = 19.10, SD = 1.32; 70.00\%$ female). As will be described
below, 23 participants completed the controllability judgment task. Based on the effect size of
$\eta_p^2 = .84$ for controllability ratings in the similar stimulus norming task in Frosch et al. (2015),
this sample size provided a power greater than .90. A similar sample of 26 participants took part
in the causal strength task.

Materials. Again, all the materials were drawn from Wells et al.’s (1987) single scenario
in which a person, trying to reach a store to buy a stereo system on sale, encounters a causal
chain of delays, ultimately resulting in the person arriving to the store after the last stereo was
sold. Eight unique causal events (i.e., delays) were assembled in keeping with a $4 \times 2$ (event:
speeding ticket, flat tire, traffic jam, senior citizens) x (intended controllability: controllable,
uncontrollable) design. The speeding ticket-controllable, flat tire-uncontrollable, traffic jam-
uncontrollable, senior citizens-uncontrollable events corresponded to the original descriptions
from Wells et al. (1987). The causal relationships between the four events in the original order
were as follows: Speeding ticket $\rightarrow$ flat tire, flat tire $\rightarrow$ traffic jam, traffic jam $\rightarrow$ senior citizens,
senior citizens $\rightarrow$ speeding ticket. To increase the likelihood of identifying stimuli that could be
balanced for controllability (of the events) and causal strengths (of the causal relationships
between events), I generated additional materials such that the causal relationship was described
in the “reverse” order as follows: Senior citizens → traffic jam, traffic jam → flat tire, flat tire → speeding ticket, speeding ticket → senior citizens.

As discussed earlier, my intent was to be able to construct two-event scenarios (i.e., three-step causal chains with a root cause, an intermediate cause, and the outcome). Each root cause was a delay event, each intermediate cause was another delay event, and the outcome was always arriving to the store after the last stereo had been sold. I will refer to the causal relationship between the root and intermediate causes of each scenario as “Link 1” and the causal relationship between the intermediate cause and arriving to the store late as “Link 2.” Thirty-two unique Link 1s resulted from a 4 (causal relationship: speeding ticket → flat tire, flat tire → traffic jam, traffic jam → senior citizens, senior citizens → speeding ticket) x 2 (order: original “forward” order, reverse order) x 2 (intended controllability of the first event: controllable, uncontrollable) x 2 (intended controllability of the second event: controllable, uncontrollable) design. In addition, 8 Link 2s were formed according to a 4 (causal relationship: speeding ticket → too late, flat tire → too late, traffic jam → too late, senior citizens → too late) x 2 (intended controllability of the first event: controllable, uncontrollable) design, wherein the outcome was always arriving to the store after the last stereo had been sold. In total, therefore, 40 unique causal relationships were constructed.

**Design.** Participants were randomly assigned to complete either a causal strength judgment task (n = 23) or a controllability judgment task (n = 26). In the causal strength task, whether people made judgments about causal relationships in the “forward” order followed by judgments about causal relationships in the “reverse” order, or vice versa, was counterbalanced. In the controllability task, whether people rated the controllability of root causes followed by the controllability of intermediate causes, or vice versa, was also counterbalanced.
Procedure. Participants in the causal strength task were presented with two-event causal relationships one at a time and were asked to rate the degree to which they believed the first event described was the cause of the second event on a 1-9 scale (1 = not at all the cause; 9 = completely the cause). In the controllability task, single events were presented one at a time to participants, who rated the degree to which they believed the event would be under their control if it happened to them on a 1-9 scale (1 = not at all controllable, 9 = completely controllable).

Results

I collapsed across counterbalancing order in reporting all of the analyses for the following reasons: First, there was no main effect of counterbalancing order; nor was there an interaction of counterbalancing order and any within-subjects factor for any of the analyses of causal strength ratings. Second, for controllability judgments, there was no interaction of counterbalancing order and any within-subjects factor; nor was there a main effect of counterbalancing order except for the analyses comprising only the senior citizens and speeding ticket events, \( F(1, 21) = 4.48, p = .046 \).

Original Wells et al. (1987) and Segura et al. (2002) materials.

Causal strength ratings. A 4 (causal event: speeding ticket, flat tire, traffic jam, senior citizens) x 2 (causal status: root cause, intermediate cause) repeated measures ANOVA was conducted on the causal strength ratings for the original stimulus materials used in Wells et al. (1987) and Segura et al. (2002). “Causal event” refers to which of the four events acted as the deeper cause in the cause-effect relationship being judged. “Causal status” refers to whether that event acted as a root cause or intermediate cause in the overall scenario from which it was taken. There was a main effect of causal event, \( F(3, 75) = 6.34, p = .001, \eta^2_p = .20, \text{BF}_{\text{incl}} = 27.311 \). Six Bonferroni-corrected paired comparisons (\( \alpha = 0.0083 \)) were performed to further examine the
main effect. These analyses revealed that the traffic jam was rated to be a stronger cause overall ($M = 6.60, SD = 1.87$) than the senior citizens crossing the road ($M = 5.23, SD = 1.68, t[25] = 4.00, p < .001, d = 0.79$).

There was also a main effect of causal status, $F(1, 75) = 20.28, p < .001, \eta^2_p = .45, BF_{incl} > 1000$. Intermediate events were rated as having more strongly caused arriving to the store late ($M = 6.59, SD = 1.88$) compared to how strongly the root causes were thought to have caused the intermediate causes ($M = 5.31, SD = 1.51$). There was no support for an interaction of causal event and causal status, $F(3, 75) = .797, p = .499, \eta^2_p = .03, BF_{incl} = .458$.

**Controllability judgments.** As with the causal strength ratings, the controllability judgments for the original stimulus materials used in Wells et al. (1987) and Segura et al. (2002) were subjected to a 4 (causal event: speeding ticket, flat tire, traffic jam, senior citizens) x 2 (causal status: root cause, intermediate cause) repeated measures ANOVA. Once again, there was a main effect of causal event, $F(3, 66) = 43.41, p < .001, \eta^2_p = .66, BF_{incl} > 1000$. Six Bonferroni-corrected paired comparisons ($\alpha = 0.0083$) showed that receiving a speeding ticket was rated as more controllable ($M = 8.20, SD = 1.18$) than the flat tire ($M = 3.93, SD = 2.22; t[22] = 8.34, p < .001, d = 1.74$), traffic jam, ($M = 3.13, SD = 2.21; t[22] = 10.34, p < .001, d = 1.91$), or senior citizens ($M = 3.50, SD = 1.91, t[22] = 10.12, p < .001, d = 2.11$) events, respectively. There was no main effect of causal status or interaction of causal event and causal status (all $Fs < 1, ps > .540, \eta^2_p < .04, BF_{incl} < .125$).

**Summary.** When analyzing the causal strength ratings and controllability judgments for the original stimulus materials used in Wells et al. (1987) and Segura et al. (2002), it becomes evident that there are significant differences between the four events in the degree to which the events are judged to have a causal impact on subsequent events as well as in the perceived
controllability of events. Segura et al. (2002)’s finding that people prefer to counterfactually “undo” root over intermediate causes in a chain of events may or may not be due to the causal status of the event as they suggested. One alternative possibility is that people’s counterfactual reasoning in that study was influenced by both causal status and controllability, separately and/or interactively. Segura et al.’s (2002) work alone cannot address this possibility, as their stimulus materials did not control for judgments of the strengths of the causal relationships between events or for the controllability of events. For instance, the strongest causal relationship was the traffic jam causing the person to have to wait for the senior citizens to cross the street, but these two events were also judged to be uncontrollable relative to the speeding ticket event. Additional work would be needed to more clearly understand patterns of counterfactual thinking about causal chains of events.

All materials. In the following analyses, I explored whether the aggregate set of all new stimulus materials that were created for this study were balanced for causal strength ratings between events, and whether people judged the controllability of the events in keeping with my intent in writing them. It was possible (and, in fact, likely) that only a subset of the materials would meet these criteria. Accordingly, in this section, I labeled subsets of the materials by order (i.e., the “forward” or “reverse” ordering) and by causal event, systematically examining selected subsets in an attempt to identify a subset that would meet the criteria. Again, the overarching goal of the current analyses was to identify three-step scenarios with equal causal strengths whose controllability ratings reflected the intended controllability of the causal events. Such materials could then be used in subsequent experiments (i.e., Experiments 4-6) to test for an interaction of controllability and causal status on counterfactual thinking, while controlling
for the previous confounds just identified in the original materials used by Wells et al. (1987) and Segura et al. (2002).

**Causal strength ratings.** Causal strength ratings for all causal pairs were entered into a 2 (causal status: root cause, intermediate cause) x 2 (intended controllability: controllable, uncontrollable) x 2 (order: forward, reverse) x 4 (causal event: speeding ticket, flat tire, traffic jam, senior citizens) repeated-measures ANOVA. I found a main effect of causal status on causal strength ratings, $F(1, 25) = 25.90, p < .001$, $\eta^2_p = .51$, $B_{F_{incl}} > 1000$. Intermediate causes were rated as a stronger cause ($M = 6.31$, $SD = 1.88$) compared to root causes ($M = 4.92$; $SD = 1.22$). Thus, critically, the event pairs for the scenarios I had created for possible use in Experiments 4-6 were not equated for causal status. This main effect was qualified by a three-way interaction of causal status, order, and causal event, $F(3, 75) = 8.03, p < .001$, $\eta^2_p = .24$, $B_{F_{incl}} = .806$, indicating some potential variability by item. Thus, it seemed that examining only a subset of causal events might provide an avenue to identify a balanced set of materials (although, as will be seen in the next section, this turned out not to be possible, as there were no combinations that would also balance controllability).

The analysis also revealed a main effect of intended controllability, $F(1, 25) = 4.59, p = .042$, $\eta^2_p = .16$, $B_{F_{incl}} = 13.510$. Events intended as uncontrollable were rated as more of a cause of the next event ($M = 5.76$, $SD = 1.41$) than were controllable events ($M = 5.47$, $SD = 1.51$). This main effect was qualified by an interaction of intended controllability and causal event, $(F[3, 75] = 6.52, p < .001$, $\eta^2_p = .21$, $B_{F_{incl}} = .057)$ suggesting that it may not hold true for all the causal events. There were no other main effects or interactions that reached significance at the $\alpha = .05$ level with a corresponding $B_{F_{incl}} > 3$ (all $Fs < 1.20$, $ps > .320$, $\eta^2_p$s < .05).
In sum, it was clear that these materials could not be used as-is in Experiments 4-6, an issue that I directly addressed in Experiment 3b.

**Controllability judgments.** Similarly, the controllability ratings were entered into a 2 (causal status: root cause, intermediate cause) x 2 (intended controllability: controllable, uncontrollable) x 4 (causal event: speeding ticket, flat tire, traffic jam, senior citizens) repeated-measures ANOVA. Order (i.e., forward or reverse) could not be entered as a factor because the causal events alone, not pairs of causal events, were rated for controllability.

There was a main effect of intended controllability, $F(1, 22) = 105.90, p < .001, \eta^2 = .83, BF_{incl} > 1000$. Events intended to be controllable were rated as more controllable ($M = 6.81, SD = 1.01$) than events intended to be uncontrollable ($M = 3.55, SD = 1.27$). However, this finding was qualified by an interaction of intended controllability and causal event, suggesting variability by item, $F(3, 66) = 12.10, p < .001, \eta^2 = .36, BF_{incl} > 1000$. No additional main effects or interactions reached significance at the $\alpha = .05$ level with a corresponding $BF_{incl} > 3$ (all $F$s < 1.87, $ps > .145, \eta^2 < .09$).

Thus, although this set of materials appeared to include a possible subset where events differed in their judged controllability as intended and did not interact with other factors, there were no such sets that were also balanced for causal strength judgments.

**Discussion**

Experiment 3a provided, for the first time, information about how people judge the causal strength and controllability – two frequently studied and highly influential factors shown to influence counterfactual thinking – of the stimulus materials used in seminal past work demonstrating that people prefer to counterfactually undo root rather than intermediate causes in a causal chain of events. Most importantly, these analyses revealed that the original materials
used in studies such as Segura et al. (2002) and Wells et al. (1987) contained causal relationships judged by participants in the current experiment that differed in causal strength. Such differences were not mentioned or otherwise previously accounted for in previous work.

Interestingly, the strongest causal relationships in the seminal materials were found between the second causal event and the outcome. Significantly weaker causal strength ratings were found between the first causal event and the second causal event. As Segura et al. (2002) and Wells et al. (1987) found a counterfactual preference to undo the first causal event, causal strength ratings may not always align with the focus of people’s counterfactual thoughts. If they did, prior work would presumably have found a much stronger counterfactual preference to focus on the last causal event than the first causal event. This pattern does not undermine previous findings suggesting that causal strength is a cue (of some kind) to counterfactual thought. Nevertheless, it is a noteworthy finding to consider when interpreting the results of past work.

The causal status of events in past work also appears to have been confounded, to some degree, with perceived controllability. Overall, participants in Segura et al. (2002) reasoned about how a controllable event causes an uncontrollable event (e.g., the speeding ticket leading to the flat tire), how an uncontrollable event causes another uncontrollable event (e.g., the flat tire leading to the traffic jam as well as the traffic jam causing the delay by the senior citizens), and also how an uncontrollable event causes a controllable event (e.g., the senior citizen delay leading to the speeding ticket). Yet no participants reasoned about the fourth possible arrangement: a controllable event causing another controllable event. In both the two- and four-event conditions in their study, participants reasoned about how an uncontrollable event caused another uncontrollable event twice as often as they reasoned about a controllable cause/uncontrollable effect or an uncontrollable cause/controllable effect. Furthermore, the root
causes were uncontrollable events three times as often as they were controllable. All of these
imbalances, revealed in Experiment 3a, may have inadvertently introduced significant confounds
into their study design.

**Experiment 3b**

The results of Experiment 3a showed that the stimulus materials used by Segura et al.
(2002) and Wells et al. (1987) cannot readily be used to cleanly test for an interaction of causal
status and controllability. It also showed that an expanded set of materials based closely on
those original materials were unbalanced for causal status (though not for controllability), and
due to the structure and content of the original materials, it appeared that this problem would not
be easily rectified. Thus, in Experiment 3b, I turned to the other common method of creating
materials to examine the influence of causal status and/or controllability on counterfactual
reasoning in the literature. This other method, found in Lagnado and Channon (2008), Frosch et
al. (2015), and Walsh and Byrne (2007), involves generating a number of different scenarios and
asking people to reason about all of them. This design has the benefit of avoiding possible
idiosyncratic effects resulting from the overall content of any one scenario or from the specific
domain. In Experiment 3b, I followed the stimulus design strategy of Frosch et al. (2015),
Lagnado and Channon (2008), and Walsh and Byrne (2007) by writing a variety of unique causal
chain scenarios. As in Experiment 3a, I measured the judged event controllability and causal
strengths of the links between the events of these scenarios, with the ultimate goal of forming a
full set of scenario materials balanced for causal strength and cleanly manipulated for
controllability to use in Experiments 4-6.
Method

Participants. Nineteen Northeastern undergraduates ($M_{age} = 19.06$, $SD = 1.35$; 73.70% female) participated. Based on the effect size of $\eta^2_p = .84$ for controllability ratings in the stimulus norming tasks used in Frosch et al. (2015) and the aggregate norming results from Experiment 3a, this sample size provided a power greater than .90.

Materials. A total of 16 scenarios were created for a stimulus norming task designed to help identify materials that were balanced for causal strength between events. Each of these scenarios consisted of a three-step causal chain resulting in a negative outcome. The scenarios largely reflected plausible everyday occurrences that I presumed participants could, with relative ease, imagine occurring to themselves. Some examples of the scenarios included traveling to a job interview, visiting a new restaurant, moving to a different apartment, and attending a birthday party. The 16 scenarios were constructed to be similar to one another in word count and overall length. Causal transition words linked the first and second events together as well as the second event to the outcome. Each of the scenarios consisted of three unique events with no overlap in content between scenarios. The temporal order of the event descriptions in each scenario concurred with their temporal order in the causal structure (e.g., no scenario first described an outcome and then subsequently described the causes). Each scenario consisted of a root cause, an intermediate cause, and an outcome; these materials were not reversible. For example, one scenario described a sign and a road barrier forcing the driver to take an alternate rather than the usual route home (the root cause), which resulted in the driver speeding through a yellow light to make up for the time lost (the intermediate cause), which resulted in an accident with another car (the negative outcome).
As in Experiments 1 and 3a, the root causes and intermediate causes were each written to convey that they were either controllable or uncontrollable by the protagonist (i.e., the study participant). For each scenario, the root cause and the intermediate cause were mismatched in intended controllability. Of the 16 scenarios tested, eight scenarios consisted of an uncontrollable root cause leading to a controllable intermediate cause resulting in a negative outcome, and the other eight scenarios consisted of a controllable root cause leading to an uncontrollable intermediate cause resulting in a negative outcome.

**Design and Procedure.** The stimulus norming task consisted of two task blocks. The order of the two task blocks was counterbalanced between participants.

In one task block, participants were informed that they would be making judgments about the causal relationship between events. As each of the 16 scenarios consisted of a root cause, an intermediate cause, and an outcome, each scenario contained two pairs of causally related events. Thus, participants read a total of 32 pairs of events. As described in the Materials section, these pairs described either the root event causing the intermediate event (i.e., Link 1) or the intermediate event causing the negative outcome (i.e., Link 2). Additionally, the first event in each link was either controllable or uncontrollable. The first event in each pair was always highlighted in yellow and the second event was always highlighted in green. The 32 pairs were presented sequentially on separate screens in randomized order for each participant. After reading each pair of events, participants answered the question, “To what degree is the event highlighted in YELLOW the cause of the event highlighted in GREEN?” on a 9-point Likert scale (1 = not at all the cause, 9 = completely the cause).

In the other task block, participants were informed that they would be making judgments of controllability. Each participant viewed the 32 separate events that corresponded to the
controllable and uncontrollable events from the 16 scenarios. Each event was highlighted in blue. The 32 events were presented in randomized order for each participant. Participants were asked to read each event description while imagining themselves experiencing the event. Then, they answered the question, “How controllable (in terms of being within YOUR control) is the event in BLUE?” on a 9-point Likert scale (1 = not at all controllable; 9 = completely controllable). Thus, this task was designed to check whether the events in the scenarios that I had intended to be viewed as either controllable or uncontrollable were in fact reliably judged as such.

**Results and Discussion**

A balanced set of materials would meet the following criteria: (1) For causal strength ratings, there should be no main effect of causal link, no main effect of intended controllability, and no interaction of causal link and controllability; (2) for controllability ratings, there should be a main effect of controllability such that people judged the controllability of events as I intended, and there should be no main effect of causal status or interaction of controllability and causal status.

**All 16 Scenarios.** First, I analyzed the causal status and controllability ratings for all 16 scenarios. For the causal strength ratings, the causal links between the root causes and the intermediate causes was link 1 and the causal links between the intermediate causes and the negative outcomes was link 2. The causal strength ratings, collapsed across scenarios, were entered into a 2 (causal link: link 1, link 2) x 2 (intended controllability: controllable, uncontrollable) repeated-measures ANOVA. There was a main effect of causal link, $F(1, 18) = 10.56, p = .004, \eta_p^2 = .37, BF_{incl} = 31.020$. Events in link 2 had higher causal strength ratings ($M = 6.72, SD = 1.10$) than events in link 1 ($M = 6.13, SD = .66$). Thus, the full set of 16 scenarios
was not balanced for the causal strength of link 1 versus link 2. There was also a main effect of intended controllability, \( F(1, 18) = 29.52, p < .001, \eta_p^2 = .62, \text{BF}_{\text{incl}} > 1000 \). Events intended to be uncontrollable had higher causal strength ratings \( (M = 6.82, SD = .76) \) than events intended to be controllable \( (M = 6.02, SD = .98) \). No other main effects or interactions reached significance at both the \( \alpha = .05 \) level and \( \text{BF}_{\text{incl}} > 3 \).

I next analyzed the controllability ratings for all 16 scenarios. The controllability ratings, collapsed across scenarios, were then subjected to a 2 (causal status: root, intermediate) x 2 (intended controllability: controllable, uncontrollable) repeated-measures ANOVA. There was a main effect of intended controllability, \( F(1, 18) = 180.85, p < .001, \eta_p^2 = .91, \text{BF}_{\text{incl}} > 1000 \). Events intended to be controllable were rated as more controllable \( (M = 7.84, SD = 1.02) \) compared to events intended to be judged as uncontrollable \( (M = 3.42, SD = 0.66) \). There was also a main effect of causal status, \( F(1, 18) = 25.52, p < .001, \eta_p^2 = .59, \text{BF}_{\text{incl}} > 1000 \). Intermediate causes were rated as more controllable \( (M = 5.95, SD = 0.64) \) than root causes \( (M = 5.31, SD = 0.44) \). There was also an interaction of causal status and intended controllability, \( F(1, 19) = 38.21, p < .001, \eta_p^2 = .68, \text{BF}_{\text{incl}} = 853.300 \). In sum, the results of the causal strength ratings and controllability ratings for all 16 scenarios revealed that this set of materials in its entirety are not adequately balanced for intended controllability across causal positions (i.e., root and intermediate causes) or for equating causal strength between link 1 and link 2.

To find a set of materials that would meet these criteria, I tested selected subsets of the stimulus materials that comprised fewer than the 16 total scenarios created (e.g., seven scenarios from the uncontrollable root cause set and seven scenarios from the controllable root cause set, etc). Appendix A contains a full description of the process used to test and identify balanced
subsets of stimulus materials for use in Experiments 4–6. I will now describe the testing results of the set of stimulus materials that was ultimately selected for this purpose.

**Balanced Subset.** Six scenarios (three in the uncontrollable root cause/controllable intermediate cause condition and three in the controllable root cause/uncontrollable intermediate cause condition) were identified (see Appendix A) and were subjected to the following analyses to verify that causal strength ratings within each scenario did not differ and that participants rated the controllability of events in line with my intent.

I collapsed causal strength ratings for these causal links across scenarios and entered them into a 2 (causal link: link 1, link 2) x 2 (intended controllability: controllable, uncontrollable) repeated-measures ANOVA. Figure 4 shows the causal strength ratings by causal link and intended controllability. There was no main effect of causal link, $F(1, 18) = .675, p = .422, \eta^2_p = .04, \text{BF}_{\text{incl}} = .239$. Thus, as intended, link 1 causal strength ratings ($M = 5.81, SD = .95$) did not differ from link 2 ratings ($M = 6.06, SD = 1.31$). There was no effect of controllability ($F[1, 18] = .019, p = .893, \eta^2_p < .01, \text{BF}_{\text{incl}} = .177$), or interaction of causal status and controllability ($F[1, 18] = .059, p = .811, \eta^2_p < .01, \text{BF}_{\text{incl}} = .059$) on participants’ causal strength ratings.

For controllability ratings, I collapsed ratings across scenario, their intended controllability, and whether they appeared as a root cause or an intermediate cause in the scenario. Controllability ratings were analyzed with a 2 (causal status: root, intermediate) x 2 (intended controllability: controllable, uncontrollable) repeated-measures ANOVA. Figure 5 shows the controllability ratings by causal status and intended controllability. There was a main effect of intended controllability, $F(1, 18) = 125.87, p < .001, \eta^2_p = .88, \text{BF}_{\text{incl}} > 1000$. The events intended to be controllable were rated as more controllable ($M = 7.78, SD = 1.11$) than
events intended to be uncontrollable ($M = 3.48, SD = .96$). There was no main effect of causal status, $F(1, 18) = .09, p = .772, \eta^2_p < .01, BF_{incl} = .290$. Finally, there was no interaction of causal status and controllability, $F(1, 18) = 2.81, p = .111, \eta^2_p = .14, BF_{incl} = .616$.

In Experiment 3b, I was thus able to create and identify a set of scenarios for which the events that I intended to be judged as controllable were rated as more controllable than the events that were intended to be judged as uncontrollable. Additionally, I found no difference in causal strength ratings between causal links that resulted from the crossing of causal status and controllability in the materials. Moreover, there was no interaction of causal status and controllability for either the causal strength ratings or controllability ratings. These balanced stimulus materials were used in Experiments 4-6.

**Experiment 4**

Experiment 4 was conducted using the new set of stimulus materials developed in Experiment 3b to examine whether causal status and controllability have an interactive influence on counterfactual thinking. To recap, I hypothesized that causal status and controllability may interactively influence counterfactual thinking. Frosch et al. (2015) showed that when an enabler and cause were mismatched in their judged controllability, people were more likely to generate a counterfactual about the controllable event compared to the uncontrollable event overall, but that when controllability was held constant, people generated counterfactuals about the enabler more so than the cause. Hilton et al. (2010) found support for an interaction of causal status and the voluntariness of an event on causal reasoning. To the degree that the effects of intentionality are similar to those of controllability, and to the extent that causal judgments align with counterfactual ones (Mandel, 2003a), event controllability may plausibly be hypothesized to interact with causal status in people’s counterfactual judgments about causal chains.
Accordingly, in Experiment 4 I asked people to generate counterfactuals about first-person, hypothetical scenarios that were characterized by a causal chain structure.

Using the materials normed in Experiment 3b, I could also uncover whether the main effect of causal status demonstrated in prior work (Segura et al. 2002; Wells et al., 1987) continues to emerge using a set of stimuli controlled for causal status, causal strength, and judged controllability. Given that Experiment 3a demonstrated that there appear to have been some imbalances inherent in the seminal stimulus materials, this was an important test to conduct with the better-controlled materials developed in Experiment 3b.

Method

Participants. I recruited 80 Northeastern University undergraduates to participate in this study (M_{age} = 19.06, SD = 1.14; 66.25% female). This sample size is in keeping with those used in prior studies indirectly supporting our hypothesized interaction (e.g., Frosch et al., 2015; Hilton et al., 2010) with effect sizes of \( \eta^2_p = .23 \) and .06, respectively.

Materials. The six scenarios that were balanced for causal strength ratings and event controllability in Experiment 3b were used in this study. Again, each of these scenarios described a causal chain sequence of two events leading to an outcome and written from the first-person perspective. Three of these scenarios consisted of an uncontrollable event causing a controllable event, which in turn caused a negative outcome. The other three scenarios consisted of a controllable event causing an uncontrollable event, which in turn caused a negative outcome.

Design. A within-subjects design was used for the main experiment, such that controllability and causal status were manipulated within each scenario as described above and presented to all participants. Participants saw all six new causal chain scenarios, one at time, in randomized order.
Procedure. All participants completed the study online using the Qualtrics software survey tool (Qualtrics, Provo, UT). For each scenario, participants were instructed to read it and to imagine themselves experiencing the scenario. Then, they were asked to complete the sentence, “Things could have turned out differently, if only…” and were given an open box to type their response. One scenario and response box were presented on the screen at a time.

After reading all six scenarios, participants completed a reading comprehension question in which they were asked to identify which of five events was not part of any of the prior scenarios. Participants then completed demographic questions and were informed about the purpose of the study.

Results

One participant (1.3% of the recruited sample) was excluded for incorrectly answering the reading comprehension question at the end of the study. The resulting analyses reported below are for the 79 remaining participants.

As in Experiment 1, two independent coders classified each participant-generated counterfactual statement as referring to the uncontrollable part of the scenario, the controllable part, both, or other. The two coders agreed on the classification of 84.78% of all counterfactuals generated and resolved disagreements through discussion. Table 3 displays the percentage of participants who generated a counterfactual by coder classification of the focus of the counterfactual and by scenario. For each participant, I calculated the percentage of all uncontrollable-first scenarios in which the counterfactual referred to only the uncontrollable event in the scenario when that event was the root cause, and the percentage of all controllable-first scenarios in which the counterfactual referred only to the controllable event in the scenario when that event was the root cause, the percentage of all controllable-first scenarios in which the
counterfactual referred to only the uncontrollable event in the scenario when that event was the intermediate cause, and the percentage of all uncontrollable-first scenarios in which the counterfactual referred only to the controllable event in the scenario when that event was the intermediate cause.

Frequentist and Bayesian analyses are summarized in Table 4. A 2 (causal status: root cause, intermediate cause) x 2 (controllability: uncontrollable, controllable) within-subjects repeated measures ANOVA conducted on the above percentages revealed only a main effect of causal status, $F(1, 79) = 94.41, p < .001, \eta_p^2 = .54, \text{BF}_{\text{incl}} > 1000$. People generated counterfactuals about intermediate causes in a higher percentage of scenarios ($M = 37.50\%, SD = 21.33$) compared to root causes ($M = 8.33\%, SD = 14.40$), an effect opposite to that found in past work. Table 3 shows that this preference was consistent across all six scenarios. There was no effect of event controllability ($F[1, 79] = .27, p = .607, \eta_p^2 < .01, \text{BF}_{\text{incl}} = .117$) and no interaction of causal status and controllability, $F(1, 79) = 1.18, p = .280, \eta_p^2 = .02, \text{BF}_{\text{incl}} = .120$. Figure 6 displays the percentage of scenarios for which people generated counterfactuals for events of each combination of causal status and controllability.

**Discussion**

In Experiment 4 I tested, for the first time, whether there was an interaction of causal status and event controllability on counterfactual thinking. I did so using stimulus materials that were verified by a separate group of participants for intended controllability and contained causal relationships that were balanced in strength. These efforts addressed concerns identified in Experiment 3a about the stimulus materials used in seminal work on counterfactual thinking.

The lack of an interaction of causal status and controllability was surprising. Past studies have provided reasons to expect it to occur: The influence of event controllability on
counterfactual thoughts was moderated by the events’ causal relationship to a negative outcome (Frosch et al., 2015), intentional human actions were preferred over natural events when making judgments of causal influence on an outcome, and this preference was stronger when human actions occurred earlier in the causal chain of events (Hilton et al., 2015). However, in Experiment 4 I found no evidence that causal status and the level of event controllability had an interactive influence on people’s counterfactual thinking, and the very low effect size of the interaction ($\eta^2 = .02$) suggests that additional power would not influence this result.

Given that Experiment 3a identified several imbalances identified in the materials used in past work on causal chains and counterfactual thought (viz. Segura et al., 2002; Wells et al., 1987), it was unknown where in a causal chain people would choose to counterfactually intervene when reasoning about chains of events that do not differ in their causal strength and that are balanced for judged controllability across all causal positions in the chain. I found that people largely preferred to generate counterfactuals about intermediate causes, and that they did so even though there was always an additional root cause event that precipitated the intermediate event in the causal chain. This is understandable in that people may believe that intervening on the intermediate event will, in general, be sufficient to prevent the outcome and can prevent the outcome regardless if the same antecedents occur again in the future or not. For example, not running a yellow light can prevent a car accident from occurring, and it can prevent an accident despite what the cause of running the yellow light may be.

In sum, Experiment 4, conducted with causal chain materials balanced for causal strength and cleanly manipulated for controllability, gave rise to three relatively surprising findings. We found the opposite of the classic causal status finding by Segura et al. (2002) and Wells et al.
(1987) for causal chains, did not replicate the original controllability finding by Frosch et al. (2015) and others, and did not find an interaction of controllability and causal status.

**Experiment 5**

Given that the use of balanced causal chain stimulus materials in Experiment 4 did not yield the same pattern of results as prior work about counterfactual reasoning on causal chains, more experiments were needed to test the reliability of the current findings. To start, one concern with the current work that needed to be addressed is that I used a counterfactual elicitation prompt in Experiment 4 (“Complete the sentence: ‘Things would have turned out differently, if only…’”) that generally led participants to only generate one counterfactual statement. Segura et al. (2002) found that the causal status of the events upon which participants chose to counterfactually intervene varied depending on whether the participant generated only one counterfactual or multiple counterfactuals. The counterfactual elicitation prompt used in their study was, “Describe how the situation could have turned out differently…” which may have prompted different responses. As the first counterfactual generated may be indicative of the primacy of people’s counterfactual beliefs (Byrne et al. 2000), only the first counterfactuals were analyzed in Segura et al. (2002). Segura et al. (2002) found that people were more likely to generate a counterfactual about intermediate rather than root causes if they only generated a single counterfactual statement (as in the current Experiment 4). However, when people generated several counterfactuals, the first counterfactual they came up with referred to the root cause more than the subsequent causes.

In order to see whether the preference to generate counterfactual thoughts about intermediate rather than root causes was an artifact of a less open-ended counterfactual elicitation prompt, Experiment 5 was conducted to determine whether the results from
Experiment 4 would be replicated or changed when people’s counterfactual thoughts were prompted by a different, more open-ended prompt as used by Segura et al. (2002).

**Method**

**Participants.** I recruited 160 Northeastern University undergraduates ($M = 18.77$, $SD = 1.19$; 77.40% female) to take part in the study for partial course credit. This total sample size was 11 times the sample needed to detect the main effect of causal status based on the effect size of from Experiment 4 ($\eta^2_p = .54$), aligning with scholarship on sufficiently powering tests of interactions (Gelman, 2018).

**Materials.** The six scenarios used in the main task of Experiment 4 were used in the current experiment.

**Design and Procedure.** As in Experiment 4, a within-subjects design was used for the main experiment, such that controllability was manipulated within each scenario and presented to all participants. I also manipulated the counterfactual elicitation prompt as a between-subjects variable. After all participants read and imagined themselves experiencing the scenario described, half the participants were presented with the same sentence-completion counterfactual prompt as in Experiment 1 and Experiment 4: “Please complete the following sentence: ‘Things would have turned out differently, if only…’.” The remaining participants were presented with the counterfactual prompt used in Byrne et al. (2000) and Segura et al. (2002): “In the box below, please describe how the situation could have turned out differently.” Again, this more open-ended prompt was included because it was more likely to elicit multiple counterfactuals compared to the other prompt used in Experiments 1 and 4, which generally lead participants to generate a single counterfactual. Although the two verbal prompts differed, the space to type responses was the same size for all participants. A total of 81 people were assigned to the
sentence-completion prompt condition and 79 people were assigned to the open-ended prompt condition.

All participants were then asked the same reading comprehension check question and demographics questions as in Experiment 4.

Results

All 160 participants answered the reading comprehension check question correctly, so all data sets were included in the analyses. The same two independent coders used in Experiment 4 (who were blind to the hypothesis of the study and the participants’ experimental conditions) classified the first counterfactual statement that the participants generated as referring to the uncontrollable part of the scenario, the controllable part, both, or other. Table 5 displays the percentage of all participants whose first generated counterfactual focused on the root cause, intermediate cause, both, and other per scenario in the sentence-completion condition. Table 6 displays the same percentages for participants in the open-ended prompt condition. In line with the intent of my counterfactual elicitation prompts, people’s written responses in the open-ended prompt condition contained multiple counterfactuals more frequently than did responses in the sentence-completion prompt condition ($\chi^2 [1, N = 984] = 35.97, p < .001$). The two coders agreed on 87.24% of responses in the sentence-completion prompt condition and 73.42% of responses in the open-ended prompt condition. For each participant, I calculated the percentage of all uncontrollable-first scenarios in which the first counterfactual generated referred to only the uncontrollable event in the scenario when that event was the root cause, and the percentage of all controllable-first scenarios in which the counterfactual referred only to the controllable event in the scenario when that event was the root cause, the percentage of all controllable-first scenarios in which the counterfactual referred to only the uncontrollable event in the scenario
when that event was the intermediate cause, and the percentage of all uncontrollable-first
scenarios in which the counterfactual referred only to the controllable event in the scenario when
that event was the intermediate cause.

Frequentist and Bayesian analyses are summarized in Table 7. A 2 (causal status: root
cause, intermediate cause) x 2 (controllability: uncontrollable, controllable) x 2 (prompt:
sentence completion, open ended) mixed-factors repeated measures ANOVA was conducted on
the above percentages. There was once again a main effect of causal status, $F(1, 158) = 10.39, p
< .001, \eta^2_p = .35, \text{BF}_{\text{incl}} > 1000$. As in Experiment 4, people generated counterfactuals about
intermediate causes in a higher percentage of scenarios ($M = 41.09\%, SD = 23.53\%$) compared
to root causes ($M = 15.61\%, SD = 20.11\%$) and the direction of means was consistent across all
six scenarios for participants in both prompt conditions (see Tables 5 and 6, respectively). There
was no effect of controllability ($F[1, 158] = .08, p = .784, \eta^2_p < .01, \text{BF}_{\text{incl}} = .073$) or interaction
of causal status and controllability ($F[1, 158] = 1.19, p = .278, \eta^2_p = .02, \text{BF}_{\text{incl}} = .056$). These
findings fully replicate the findings of Experiment 4.

There was an interaction of controllability and prompt, $F(1, 158) = 6.51, p = .012, \eta^2_p = .04, \text{BF}_{\text{incl}} = .219$. I ran four Bonferroni-corrected paired comparisons ($\alpha = .0125$) to better
understand this interaction. First, people receiving the open-ended prompt generated
counterfactuals about controllable events in a higher percentage of scenarios ($M = 32.07\%, SD =
19.75\%$) than did participants receiving the sentence-completion prompt ($M = 24.07\%, SD =
18.45\%$), $t(158) = 2.65, p = .009, d = .42$. There was no difference between prompts for the
percentage of scenarios wherein people generated a counterfactual about uncontrollable events,
$t(158) = 2.65, p = .403, d = .13$. Taken together, these comparisons suggest that having the
flexibility to generate multiple counterfactuals may lead people to think more about what could
actually have been done to change the situation as it unfolded, an unexpected finding. Second, I
did not find an effect (when corrected for familywise error) of controllability in the sentence-
completion prompt condition alone, \( t(80) = 2.03, p = .046, d = .23 \), or in the open-ended prompt
condition alone, \( t(78) = 1.59, p = .117, d = .18 \). Thus, the lack of overall preference for
controllable over uncontrollable events was consistent across two counterfactual prompt
manipulations.

No other main effect or interactions reached significance at the \( \alpha = .05 \) level. Figure 7
displays the percentage of scenarios for which people generated counterfactuals for events of
each combination of causal status and controllability, collapsed across prompt.

**Discussion**

Experiment 5 replicated the results of Experiment 4 using two different methods to elicit
counterfactual statements by participants. In Experiment 5, I used the sentence-completion
prompt (Frosch et al., 2015) as well as the legacy open-ended response prompt used in prior
work (Segura et al., 2002; Byrne et al. 2000) to see whether the open-endedness of the prompt
influenced the focus of people’s primary counterfactual. Once again, I found that people
preferred to generate counterfactuals about intermediate causes rather than root causes. While
this was initially a surprising finding in Experiment 4, this pattern has now emerged across two
experiments, and has appeared to be robust to different types of prompts commonly used to elicit
people’s counterfactual thoughts. Thus far, my data have consistently indicated that when
people reason about causal chains that are balanced for their level of controllability and the
causal strength between the events, people prefer to counterfactually intervene on intermediate
causes rather than root causes. Thus, when correcting for multiple imbalances identified in
Experiment 3a in past work on the influence of causal status on counterfactual thought (e.g.,
Segura et al., 2002; Wells et al., 1987), the preference to “undo” root causes over subsequent causes in the chain is no longer found to occur. These findings advance new knowledge concerning the influence of causal status on counterfactual thought.

I found no interaction of counterfactual elicitation prompt and causal status in Experiment 5, thus replicating the results of Experiment 4 and extending the same findings across two methods used in counterfactual reasoning tasks to elicit counterfactual thoughts from reasoners. Thus, once again, although prior work has provided reasons to suspect an interaction of causal status and controllability on counterfactual thought (e.g., Frosch et al., 2015; Hilton et al., 2010), the current experiments have not found consistent evidence for such an interaction.

Together, these results support the interpretation that the counterfactual preference for intermediate rather than root causes is a consistent finding that occurs when stimulus materials are balanced for causal strength judgments and are counterbalanced for event controllability across causal status positions. In addition, just as in Experiment 4, I found evidence that controllability had no influence on counterfactual thought. Prior work indicating that the judged controllability of events influences the focus of counterfactual thinking has employed scenarios that were not balanced for causal strength between events (Frosch et al., 2015) or by having participants each reason about a unique, personally-relevant event (Mandel, 2003b; Markman & Miller, 2006). It has been unknown whether event controllability influences counterfactual thought when people reason about well-controlled scenarios wherein events do not differ in their causal strength. Thus far, Experiments 4 and 5 have provided empirical evidence that counterfactual thought is not influenced by the controllability of events when the causal links between them have been reliably verified to not differ in their judged causal strength. As will
later be seen, Experiment 7 will further contribute to our understanding of the influence of controllability on counterfactual thought.

Experiment 6

The results of Experiment 5 replicated the findings from Experiment 4, and the counterfactual preference to “undo” intermediate rather than root events was replicated and was not found to be an artifact of the way in which people’s counterfactuals were elicited. A remaining concern in interpreting the results of Experiments 4 and 5 (and 1), however, is that the coding decisions of the external raters may not have reflected the actual counterfactual intention of the participants. Given the challenges inherent in interpreting open-ended statements, which varied in their level of ambiguity, the raters’ coding of the counterfactuals may have diverged from what participants themselves would have to say about their responses. It is certainly possible that the accuracy of coding may have depended upon the clarity of participants’ writing and/or writing ability more generally. To address the possibility that the results were dependent on the decisions made by coders who may or may not have been able to correctly discern participants’ intentions in writing about their counterfactual choices, Experiment 6 was conducted. In Experiment 6, I asked whether the results from Experiment 4 would continue to emerge when people coded their own counterfactual responses.

Method

Participants. I recruited 83 Northeastern University undergraduates (\(M_{\text{age}} = 18.81, \ SD = 0.99, 67.50\% \text{ female}\)) to take part in this study for partial course credit.

Materials and Procedure. We used the same six scenarios from Experiments 4 and 5 in this study. Each participant was presented with the six scenarios in randomized order, and was asked to generate a counterfactual about the scenario. Because Experiment 5 showed that the
results did not depend on the type of prompt used, I used only the single-counterfactual prompt used in Experiment 4 (which was the single-counterfactual condition in Experiment 5: “Complete the sentence ‘Things would have turned out differently, if only…””).

After generating counterfactuals for all the scenarios, participants then coded their own written counterfactuals in the same way that the independent coders from Experiment 4 and 5 coded the participant-generated counterfactuals. Participants were presented with the scenarios again, along with their earlier written counterfactual responses, one at a time. The root cause of the scenario was now highlighted in yellow and the intermediate cause of the scenario was highlighted in green. The highlighted contents in the scenario corresponded to the highlighted sections used in the stimulus norming task used in Experiment 3b. Participants were asked to select the statement that best reflected their own written counterfactual response. The four options were whether their written responses (1) referred primarily to the event highlighted in yellow (i.e., the root cause), (2) referred primarily to the event highlighted in green (i.e., the intermediate cause), (3) referred equally to the yellow and green events, or (4) referred primarily to events not highlighted in the scenario. Finally, participants answered the same reading comprehension check question as used in Experiment 4 and 5 and a demographic questionnaire.

Results

One participant (1.2% of the total sample) was excluded from the following analyses for incorrectly answering the reading comprehension question. Table 8 displays the percentage of all participants who generated a counterfactual about the root cause, intermediate cause, both, and other across scenarios in the sentence-completion condition.

Frequentist and Bayesian analyses are summarized in Table 9. A 2 (causal status: root cause, intermediate cause) x 2 (controllability: uncontrollable, controllable) repeated-measures
ANOVA was conducted on the same counterfactual percentages. Figure 8 displays the percentage of scenarios for which people generated counterfactuals for events of each combination of causal status and controllability. The results revealed, once again, a main effect of causal status, $F(1, 81) = 91.45, p < .001, \eta^2_p = .53, \text{BF}_{\text{incl}} > 1000$. People classified a larger percentage of their own counterfactual statements as referring to intermediate causes ($M = 60.77\%, SD = 24.45$) than root causes ($M = 17.02\%, SD = 20.83$). Once again, this pattern was consistent across all six scenarios about which people reasoned (see Table 5). There was no effect of controllability ($F(1, 81) = .56, p = .46, \eta^2_p = .01, \text{BF}_{\text{incl}} = .140$) and no interaction of causal status and controllability, $F(1, 81) = .02, p = .901, \eta^2_p < .01, \text{BF}_{\text{incl}} = .102$.

Discussion

Experiment 6 was conducted to test whether the results of Experiments 4 and 5 would be replicated when the participants themselves, rather than independent coders, classified the counterfactuals that they generated as referring to the root or intermediate cause. The participants’ coding data once again revealed a main effect of causal status, such that more counterfactuals “undid” the intermediate rather than the root cause. Participants’ own coding responses did not suggest an effect of event controllability or an interaction of causal status and controllability. A secondary analysis of the counterfactuals that were generated in this experiment using classifications made by independent coders is reported in Appendix C. The results of that secondary analysis are also largely in line with the pattern of results of the current Experiment 6 and the earlier Experiment 4. In sum, across Experiments 4–6, evidence for an interaction of causal status and controllability in counterfactual thinking was not found.
**Discussion of Experiments 3a – 6**

Prior work has indicated that the causal status of events influences where people choose to counterfactually intervene when imagining how a situation could have turned out differently. Specifically, Segura et al. (2002) and Wells et al. (1987) found evidence that people prefer to generate counterfactuals about root causes rather than subsequent intermediate causes when reasoning about events in a causal chain. However, as I found in Experiment 3a, their stimulus materials were not pretested for causal strength or event controllability; nor were they balanced for those factors. In Experiments 4-6, I sought to determine whether past findings documenting the effects of causal status and of controllability on counterfactuals would be replicated when the stimulus materials were well controlled and counterbalanced. The other central concern of Experiments 4 through 6 was to examine whether event controllability and causal status interactively influence the formation of counterfactual thoughts.

Experiment 3a provided evidence that the stimulus materials used in past work (i.e., the classic scenario about a shopping experiencing a series of delays on his way to a sale) have significant confounds and imbalances that make them less than ideal for testing for an interaction of causal status and controllability. I found that the different causal links in the scenario’s chain of events varied significantly in terms of their causal strength, making it more likely that people would generate counterfactuals that would affect the strongest links. I also showed that the scenario contained three relatively uncontrollable events and only one relatively controllable event, suggesting that people’s mental “undoing” would focus on the sole controllable event (Frosch et al., 2015). These confounds called into question previous arguments that people generally prefer to generate counterfactuals about root cause over intermediate cause events and that they generally prefer to undo controllable over uncontrollable events.
Given these limitations, I designed Experiment 3b to create, pretest, and identify a wide variety of scenarios that were balanced for event controllability and causal strength. In Experiment 4, I used a balanced subset of these new materials to test, for the first time, whether controllability and causal status interactively influence the generation of counterfactual thoughts. Experiment 4 found no evidence for an interaction, but did find strong evidence that people preferred to generate counterfactuals about intermediate rather than root causes, a reversal of the findings in the earlier literature. This finding was replicated in Experiment 5 using a different method of counterfactual elicitation, and was replicated once more in Experiment 6 when the participants themselves, in addition to independent coders, classified their counterfactual statements. These findings differ markedly from the pattern of counterfactual thinking about causal chains identified by Segura et al. (2002). However, they are analogous to Lagnado and Channon’s findings (2008) that people attribute greater causal responsibility to intermediate than root causes when reasoning about many everyday examples of causal chains. Thus, it may be the case that causal and counterfactual reasoning operate similarly when reasoning about everyday causal chains.

Surprisingly, using the newly pretested and balanced materials, I found no influence of event controllability on people’s counterfactual thoughts. Prior work found that event controllability influences counterfactual thought about negative events in one’s own life (Mandel, 2003b; Markman & Miller, 2006), sequences of events comprising enabling conditions and direct causes (Frosch et al., 2015), and chains of events containing human actions as well as non-human events in nature (Hilton et al., 2010). Yet, Experiments 4 through 6 were the first to show that when event controllability and causal strength were balanced across scenarios in a causal chain structure, the relative controllability of the events did not influence where in the
causal chain people preferred to counterfactually intervene. At the very least, the current work calls into question past research on the effects of causal status and controllability on counterfactual thinking. Finally, it cannot be argued that event controllability moderates the influence of causal status; I found little support in Experiments 4 through 6 that event controllability and causal status interactively influenced counterfactual thought.

**Experiment 7**

My findings thus far do not support a reliable effect of the controllability of events on counterfactual thinking. However, other prior research has found a predictive relationship between the controllability of events and changes in affect following counterfactual reasoning (Mandel, 2003b) such that individuals experience more negative affect (e.g., feeling more bad, sad, and negative) when engaging in counterfactual thinking about controllable events. As I described in the Introduction, when people generate upward counterfactuals (i.e., counterfactuals about how a better outcome could have been achieved) about events in their own lives, they usually experience increased negative affect after doing so (Mandel, 2003b; Markman & Miller, 2006; Markman & Weary, 1996). Previous research has suggested that the controllability of events may be influential in bringing about this change in affect (Mandel, 2003b).

Past studies have often used scenario-based designs to assess how counterfactual availability influences perceptions of other’s emotions (Gleicher et al. 1990; Kahneman & Miller, 1986; Landman, 1987). While these studies have provided valuable information about the relationship between counterfactual thought and affect, a richer understanding of such a relationship might be gained by also having people reason counterfactually about events in their own lives and self-reporting affect before and after doing so.
Accordingly, Experiment 7 was designed to address two questions. First, I asked if Mandel’s (2003b) findings (i.e., whether the controllability of events influences changes in affect after generating upward counterfactuals) would be replicated when people reason about negative events that really happened to them. Second, I asked to what degree the causal status of counterfactuals might also influence affect change following upward counterfactual thinking, as Experiments 4-6 showed that causal status reliably influenced counterfactual thinking.

Experiment 7 was an exploratory study conducted to address these questions. As in Experiments 1, 4, 5, and 6, I asked people to generate counterfactuals about a focal event. However, as I argued above, a different methodology was needed to more accurately elicit any potential changes in affect. Accordingly, I asked people to think counterfactually about a recent negative event in their own life in order to increase the realism of the study and increase the likelihood of evoking a relatively strong affective reaction from the participants. As the events in people’s own lives may likely be more complex than the causal chain scenarios studied in Experiments 3a through 6, I employed a causal model technique to compute the causal status of the events in the negative scenarios people described, and to define accordingly where in this causal structure people’s counterfactuals intervened. This allowed me to understand any relationship between the causal status of counterfactual thought and affect change.

Method

Participants. Based on the effect size of affect change following counterfactual thought ($\eta_p^2 = .39$) in a similar work (Markman & Miller, 2006), a sample of 75 participants was needed to achieve power greater than .80. However, given that participants generated their events in the
current study, to ensure sufficient power I recruited 102 Northeastern University undergraduate students ($M_{\text{age}} = 18.47, SD = 0.89; 74.51\%$ female) to take part for partial course credit in an introductory psychology course.

**Design and procedure.** Adapting procedures from prior research using participants’ self-generated negative events (Markman & Miller, 2006; Markman & Weary, 1996; Mandel, 2003b), I asked participants to think of, and write about, a recent negative life event they had experienced. Following this, participants were asked how negative, bad, and sad thinking about the event made them feel, and to report the amount of control they felt they had had over the event. All four questions were answered using a 9-point Likert scale (1 = *not at all*; 9 = *extremely*; Markman & Weary, 1996). After making these ratings, participants saw a new screen with the following text:

After experiencing negative outcomes like the one you just described, people often can’t help thinking, “if only…” and imagining how the outcome could have turned out better. On the following screen, list as many examples of “if only” thoughts that come to mind as you think about the negative outcome.

Below this paragraph was a blank text box in which participants typed in their counterfactual statements. Then, participants were again presented with the three affect-related questions and the question about perceived control, and were asked to answer them while reflecting on their negative life event.

**Results**

**Data preparation and description of analyses.** Of the 102 participants recruited, a total of 26 participants were removed from analyses for not following experimental instructions or for substantial coding difficulties. For instance, some participants did not adhere to the instructions
for writing about a recent negative event. Some participants wrote about a recent event, but explicitly stated that the event was not negative or distressing to them (violating criterion 3 from the instructions). Other participants wrote about events that did not happen to them, instead writing about negative events experienced by a friend or roommate (violating criterion 2). Other participants did not generate any counterfactual statements at all. Finally, some participants were removed from analyses if coders could not agree with confidence on how to record causal maps for those participants’ events, typically for reasons of incoherence in writing.

**Generation of causal maps.** Three coders independently constructed a causal map representing each participant’s written description. Specifically, for each participant, coders first read the event description and composed a list of candidate features of the event. For instance, one participant wrote, “I forgot about an upcoming chemistry test. I did not have time to study adequately so I failed my test”. Coders would create the following features from this written description: *forgot about chemistry test, did not have time to study*, and *failed test*. Coders came to a consensus decision about the inclusion of features by discussion. In general, features were created when explicitly mentioned in the participant’s written event description. That is, coders did not list features of the event that they assumed occurred. The lone exception to this rule was that the feature of *negative affect* was added to all participants’ causal maps as a feature regardless of whether the participant explicitly mentioned it or not, because participants had been instructed to write about something negative that had happened to them.

Coders also coded the list of the counterfactual thoughts that the participant wrote during the experiment. These counterfactual statements were also represented as features. For instance, if participants wrote as a counterfactual statement “If only I wasn’t so bad with scheduling and organization…” then the coders would add the feature *was bad with scheduling and*
organization. If participants wrote “If only the professor had a flat tire on his drive to campus”, the coders would add the feature *professor did not have flat tire*.

After generating the above features, the coders constructed a causal map in which all of the features were included. Using the ConceptBuilder software package (Kim & Park, 2009), the coders independently organized the event features in a causal diagram. In ConceptBuilder, each feature appears as a boxed node and users can draw causal links between these nodes using arrows. Going back to the above example, the coders would draw a causal link between *forgot about chemistry test* and *did not have time to study*. While ConceptBuilder allows users to assign causal links different weights for different causal strengths, the weight of all arrows was set to the lowest weight of 1.\(^7\) Via discussion and consensus, the three coders then arranged all of the nodes into a single causal map representing the written description of the negative life event and the listed counterfactual thoughts. Features could serve as common effects and common causes for other features. Causal cycles were also allowed in the maps. For instance, if a participant wrote about being depressed because of his poor grades and also that his depression was a cause of his poor grades, then there would be a two-way causal link between *being depressed* and *poor grades*. Features were left isolated if they were judged to be neither a cause nor an effect of any other feature. All causal maps were created by consensus.

**Causal centrality scores for features.** Causal centrality scores (Sloman, Love & Ahn, 1998) were calculated from the causal maps created by the coders using the ConceptAnalysis software package (Kim & Park, 2009), which computes eigenvector centrality for directed graphs (see Kim, Luhmann, Pierce, & Ryan, 2009). ConceptBuilder first generates an \(n \times n\) adjacency matrix where \(n\) is the number of features present in the event’s causal map. Using columns to represent causes and rows to represent effects, the adjacency matrix is created by
placing a value of 1 at the intersection of any linked cause feature with any linked effect feature. A value of 1 is also created at the intersection of any feature with itself. ConceptAnalysis employs the following equation to compute the causal centrality ratings:

\[ c_{i,t+1} = \sum_j d_{ij} c_{j,t} \]

where \( d_{ij} \) is a positive number that represents the degree to which feature \( j \) is a downstream effect of cause \( i \), and \( c_{j,t} \) is the causal centrality of feature \( j \) at time \( t \) (Kim & Park, 2009). ConceptAnalysis iteratively computes the causal centrality ratings until they reach stable values. The final causal centrality values are meaningfully interpreted in terms of their relative rank order.

**Causal status of counterfactual statements.** I used a Wilcoxon Rank Order test to compute the relative causal status of the counterfactual features as compared to the non-counterfactual event nodes in the causal map. For instance, it is possible that participants’ counterfactual features tended to have higher causal centrality (i.e., a higher causal status) than the other features. The Wilcoxon Rank Order test yielded a z-score signifying the rank-order distribution of counterfactual nodes compared to the rank-order distribution of the rest of the non-counterfactual event nodes.

In the context of the current calculations, a z-score with a large negative value would indicate that the counterfactual nodes had a higher causal centrality than the non-counterfactual nodes. That is, it would indicate that the participant generated counterfactuals for more causally central (i.e., early) features than for subsequent, causally downstream features. This would be roughly analogous to creating counterfactuals about the first or second event in a four-event causal chain structure as depicted in Figure 1, except that the current causal maps were more complex. A z-score with a large positive value would indicate that the counterfactual nodes had
a lower causal centrality compared to the non-counterfactual nodes. This might be roughly analogous to creating counterfactuals about the third or fourth event in a four-event causal chain structure (Figure 2).

The use of the Wilcoxon Rank Order z-score was also necessary because the structure of each participant’s event causal map was unique. Therefore, this computation allowed for meaningful comparisons to be made concerning the causal position of the counterfactual nodes across various maps. Since the total number of nodes in a given causal map can vary greatly, the absolute ranks for one participant’s counterfactual nodes could not otherwise be meaningfully compared to the ranks for another participant’s counterfactual nodes.

**Causal centrality and changes in affect.** Participants’ scores for the three negative affect questions were summed to yield two separate composite scores for pre- and post-counterfactual affect ratings (Cronbach’s alpha of .85 and .91, respectively). Post-counterfactual scores were subtracted from the pre-scores to determine the change in affect following counterfactual thinking. Negative change scores indicated that people felt less negative affect when reflecting on the negative event after completing the counterfactual task compared to reflecting on the negative event before the counterfactual task.

We found a significant negative association between the causal status of people’s counterfactual statements and their improvement in affect ($r = -.23, p = .045$). Specifically, the more central the causal status of people’s counterfactual nodes relative to the rest, the greater the negative change in people’s affect following the upward counterfactual generation task.

**Causal centrality and changes in affect by event type.** As discussed in the introduction to this experiment, past work addressing the functional perspective on counterfactual thinking has uncovered some differences in the counterfactual thinking of undergraduate students when
reasoning about two different types of negative events: Academic events and interpersonal events (Mandel, 2003b). These two types of events have been shown to differ in the degree to which people feel the outcome was controllable or not. Specifically, Mandel (2003b) found evidence that academic events are judged by individuals as more controllable in nature compared to interpersonal events. He explained:

For many academic setbacks, students may be aware that consensus is low: The exam that they failed, for instance, was probably passed by most of the class. Low consensus information is consistent with internal attributions (Kelley, 1967). Moreover, most students know that academic institutions set the “rules of the game” and it is their responsibility to learn how to play by them. Students therefore are likely to see their actions as figural against a background of constraints that include the professor’s actions (McGill, 1989; Miller & Gunasegaram, 1990). By contrast, people are apt to perceive interpersonal situations as being governed by principles of cooperation, fairness, and reciprocity. Unlike in academia, there is a sense in which the rules of the game ought to be informally, and concurrently, set by the various individuals involved. When interpersonal conflicts arise, people, each with their own egocentric focus, may be likely to defensively attribute responsibility and blame to others (Shaver, 1970). (Mandel, 2003b; pp. 141-142)

As it turned out, in Experiment 7, coders identified all negative events described as either academic or interpersonal events. Given Mandel’s (2003b) evidence that academic and interpersonal events differ with respect to how controllable people tend to judge them, I also conducted the analyses broken down by event type. No participants wrote solely about a negative experience that could not be understood as occurring within one context or the other.
For example, no participants wrote about a negative experience with only an object (e.g., writing solely about frustration with a defective laptop computer). Participants who did write about frustration with objects did so in the context of an interpersonal interaction (e.g., arguing with a mechanic about insufficient repairs made on an automobile). The coders identified 35 academic events and 40 interpersonal events. Table 10 displays pre- and post-counterfactual affect scores as a function of event type. A correlational analysis revealed no relationship between counterfactual causal status and affect change for academic events ($r = .00, p = .998$). However, there was a significant relationship for the interpersonal events ($r = -.47, p = .002$). This pattern was similar to the overall pattern for all events reported above, such that affect became more negative as the causal status of the counterfactuals increased. That is, the relationship between the causal status of the counterfactuals and affect change appears to have been entirely driven by the interpersonal events.

**Secondary analyses: Event type and counterfactual causal status.** I also explored whether event type predicted the causal status of the counterfactuals people generated. Indeed, there was a significant difference in relative causal status between event types. Counterfactual node z-scores were more negative for academic events ($M = -1.74, SD = 1.17$) than for interpersonal events ($M = -1.01, SD = .89$), $t(73) = -3.05, p = .003, d = .70$. In other words, people’s counterfactual statements were more causally central for academic than interpersonal events. There was also a greater percentage of academic (51.4%) events with a significant counterfactual node z-score (less than or equal to -1.96) compared to interpersonal events (10.0%). Thus, it is not just that people who wrote about academic events merely had a higher number of causally deep counterfactual statements relative to interpersonal events. The
counterfactual statements for the academic events were also significantly causally deeper (i.e., more central) compared to the interpersonal events.

**Controllability of counterfactual statements.** Two new coders who did not construct the causal maps and who were blind to the hypotheses coded each participant’s counterfactual statements for whether they were uncontrollable (e.g., “If only a blizzard had postponed the test...”) or controllable (e.g., “If only I had started studying earlier...”) by the participant, in accord with coding procedures from past work on counterfactual thinking and controllability (c.f., Mandel, 2003b; Markman & Miller, 2006; Markman & Weary, 1996).

Overall, there was no correlation between the proportion of counterfactuals generated that were judged to be controllable and change in negative affect, \( r(98) = .08, p = .457 \). However, when split by event type, the correlation was significant for academic events, \( r(35) = .35, p = .037 \). The greater the proportion of controllable counterfactuals that people generated, the more that people’s negative affect increased after upward counterfactual thinking. There was no significant relationship between these two variables for interpersonal events, \( r(37) = -.05, p = .749 \).

**Regression analyses.** I ran two multiple regression analyses to address whether higher causal status still predicts change in affect even when controlling for the number of controllable counterfactuals generated, which was found to differ significantly between domains with more controllable counterfactuals generated for academic compared to interpersonal events. The number of controllable counterfactuals and the causal status z-score were entered to predict change in affect. For academic events, the overall model was significant, \( F(2, 32) = 4.28, p = .023, R^2 = .20 \). The number of counterfactuals predicted increased negative affect (\( \beta = .69, p = .006 \)) but the causal status of the counterfactuals did not (\( \beta = .23, p = .570 \)).
The opposite pattern was found for the interpersonal events. The overall model was significant, $F(2, 37) = 5.28, p = .010, R^2 = .22$. However, for interpersonal events, higher causal status of the counterfactuals predicted increased negative affect ($\beta = -1.78, p = .003$) but not the number of controllable counterfactuals ($\beta = -.02, p = .915$). Thus, when controlling for the number of controllable counterfactuals people generated, there was still a meaningful relationship between higher causal status and subsequent change in affect.

**Secondary analyses.** I carried out two additional sets of analyses that were not the primary focus of the current experiment.

*Secondary analyses: Number of controllable and uncontrollable counterfactual statements by event type.* Past work has reported that people are likely to generate a higher ratio of counterfactuals that focus on controllable rather than uncontrollable aspects of an event (Markman & Miller, 2006). Likewise, people are more likely to generate upward counterfactual thoughts following negative outcomes that were judged to be relatively controllable in nature (e.g., academic events) than for those judge to be relatively uncontrollable (e.g., interpersonal events; Mandel, 2003b). I therefore analyzed whether such a pattern generalized to situations in which people were not directed to write about one type of event or the other. A 2 (event type: academic, interpersonal) x 2 (controllability: controllable, uncontrollable) repeated-measures ANOVA revealed a main effect of controllability, $F(1, 73) = 58.055, p < .001, \eta_p^2 = .443$, such that people writing about either event type generated more controllable ($M = 3.28, SE = .24$) than uncontrollable ($M = .86, SE = .14$) counterfactuals. There was also a main effect of event type, $F(1, 73) = 13.43, p < .001, \eta_p^2 = .16$. People generated more counterfactuals overall for academic events ($M = 2.50, SE = .17$) than for interpersonal events ($M = 1.63, SE = .16$). Finally, there was a significant interaction of event type and controllability, $F(1, 73) < .01, p = \ldots$
Bonferroni-corrected paired comparisons revealed that people generated more controllable counterfactuals for academic events ($M = 4.22, SE = .36$) than interpersonal events ($M = 2.33, SE = .35$), $t(73) = 3.92, p < .001, d = .91$. In contrast, I found no effect of event type on the number of uncontrollable counterfactuals that people generated, $t(73) = -.65, p = .520, d = -.15$.

The above patterns of results seem to indirectly suggest that undergraduates perceived the academic events as more controllable than the interpersonal events, in line with Mandel’s (2003b) suggestion. However, it was not reflected in their ratings of the events overall. A $2 \times 2$ (event type: academic, interpersonal) x 2 (control rating time: pre-, post-counterfactual task) repeated measures ANOVA revealed that whereas people generally felt the event was more controllable after completing the counterfactual thinking task than they had before ($F[1, 73] = 7.75, p = .007, \eta^2_p = .10$). This effect was not moderated by event type, as there was no interaction of event type and rating time ($F = .00, p = 1.00$); nor was there a main effect of event type ($F[1, 73] = .07, p = .796, \eta^2_p < .01$).

**Discussion**

The results of this exploratory study suggest that increased negative affect following upward counterfactual thinking may be dependent on the degree to which counterfactual thoughts generated target causally deep features of an event, and on the domain of the event under consideration. Generating more causally deep counterfactuals predicted greater increases in negative affect, but only for academic events, not for interpersonal events.

There are several limitations to the current study. First, there may be a confound of counterfactual controllability and causal status, as we were unable to systematically control either in this experiment. Second, although the domain of events was of interest, the event
domain that participants wrote about was not systematically controlled for between subjects. Participants were free to write about any negative event with no requirements with regard to domain. Thus, participants self-selected themselves into the academic and interpersonal domain groups, allowing the possibility that individual differences underlying event domain also account for changes in affect. Finally, both the causal structure and the relative controllability of the events were created and judged by external coders. As such, participants may not fully endorse the arrangement of the nodes in the causal map or the ratings of whether their counterfactuals are controllable or not. Future work will be needed to determine whether the current findings can be replicated when people themselves are able to determine the causal structure and controllability of their negative events, with domain manipulated within the experimental design.

**General Discussion**

Counterfactual thinking is a mental process whereby people mentally “undo” events that have occurred and simulate what the resulting outcome could have been, in contrast to what actually occurred (Byrne, 2016; Kahneman & Miller, 1986). People engage in this mental simulation often about events in their own life, and it typically takes the form of people generating “upward” counterfactuals whereby people imagine how a negative outcome could have turned out better. Below, I consider my three central aims in light of the results of the current work, and how these findings advance new knowledge about people’s counterfactual thought.

**Aim #1: Testing for a Correspondence Bias in Counterfactual Thought**

Experiment 1 found no evidence that point of view moderated people’s preference to counterfactually “undo” controllable rather than uncontrollable events, replicating the findings of Frosch et al. (2015) in both first- and third-person reasoners. Experiment 2 found no effect of
point of view on where in a causal system people chose to intervene. Across both perspectives, people preferred to counterfactually intervene on root-cause events over subsequent events in a causal-chain structure. However, in contrast to prior work (e.g., Segura et al., 2002), people displayed no preference for where to intervene in a conjunctive common-effect structure, and point of view did not moderate this.

Together, these two studies replicated work that demonstrates people’s preference to generate counterfactuals for controllable over uncontrollable events and that people prefer to intervene on root events in a causal chain structure when reasoning from a third-person perspective and extended these findings to reasoning from a first-person perspective. Given that point of view did not moderate the effects of controllability and causal status, I asked people to reason in the first-person perspective in the remaining studies to test for an interaction of controllability and causal status on counterfactual thought (Experiments 3a–6) and for the influence of controllability and causal status on affect change (Experiment 7).

**Aim #2: Testing for an Interaction of Event Controllability and Causal Status on Counterfactual Thought**

Experiment 3a found that the stimulus materials used in past work advancing the findings that people are most likely to counterfactually intervene on root causes in chain (Segura et al., 2002; Wells et al., 1987) had several potentially critical imbalances and confounds. For instance, uncontrollable events occurred in the root causal position three times as often as controllable events. Some events were also judged to be significantly stronger causes than others, suggesting that the links in the chain are not balanced in terms of judged causal strength. Thus, those materials could not be used in Experiments 4–6 and also called into question whether the counterfactual preference for root causes would continue to emerge once people reasoned
about balanced chains of events. Experiment 3b yielded a new set of stimulus materials that did not have these confounds (i.e., the two links in the causal chain did not differ in terms of their rated causal strength and event controllability was balanced across causal positions in the chain).

Experiment 4 then constituted my first direct experimental test of whether causal status and controllability interactively influenced counterfactual reasoning, and whether past work showing independent effects of causal status and controllability on counterfactual reasoning would be replicated using more carefully controlled materials. I found that causal status and controllability did not interact to influence people’s counterfactual thoughts. However, two other results emerged from this study. First, although I did find a main effect of causal status, the preference displayed was such that people preferred to counterfactually intervene on intermediate rather than root causes across all six scenarios that people reasoned about. Thus, the preference to undo root causes displayed by Segura et al. (2002) and Wells et al. (1987) did not continue to emerge once people reasoned about a balanced set of materials. Second, I did not find an effect of event controllability as anticipated by prior counterfactual reasoning studies (e.g., Frosch et al., 2015) and other research falling under the umbrella of the functional framework of counterfactual thinking (e.g., research supporting the view that mental simulations may be deployed as a means to lead people to identify controllable counterfactuals to achieve more optimal outcomes in the future; Epstude & Roese, 2008; Roese & Epstude, 2017). People were not, in fact, more likely to generate counterfactuals about controllable events than about uncontrollable events across the balanced set of stimulus materials.

As these results of Experiment 4 contrasted with those found in past work, I conducted additional studies to determine whether this effect would be replicated and also whether it would continue to emerge across additional means of eliciting and classifying counterfactual thoughts.
Experiment 5 was a replication of Experiment 4 using an alternate prompt to elicit counterfactual thoughts, as previously used in the counterfactual reasoning literature, to see whether a more open-ended counterfactual elicitation technique would moderate the findings from Experiment 4. Instead, I found that across two prompt conditions, that there was again no interaction of causal status and controllability. The preference to “undo” intermediate rather than root causes was found again in both conditions and there continued to be no preference for controllable over uncontrollable events across both conditions.

Experiment 6 was designed to address concerns that there may have been an information asymmetry between participants and coders in Experiments 4 and 5 that lead to coding decisions that did not reflect participants’ intentions in generating the counterfactual. To test this possibility, I had people first generate counterfactuals about the six scenarios in used in Experiments 4 and 5 and then classify them. Participants were well able to complete this task and their pattern of coding produced the same pattern of results as found in Experiments 4 and 5. There was again the preference to “undo” intermediate causes rather than root causes, no effect of event controllability, and no interaction of causal status and controllability.

Overall, Experiments 4–6 provided little evidence to support an interaction of causal status and controllability on counterfactual thought. These studies also provided consistent evidence that people prefer to generate counterfactuals about intermediate more than root causes in a causal chain of events when reasoning about a balanced set of materials. Finally, Experiments 4-6 did not provide evidence that the controllability of events influences counterfactual thinking. However, as will be described next, the much more open-ended and personalized Experiment 7 did provide evidence that people generate more counterfactuals focusing on controllable rather than uncontrollable events. Additionally, the domain of the focal
event moderated the influence of event controllability. Thus, more work is needed using self-relevant and open-ended approaches; in addition, more work is needed that uses well-controlled materials spanning a wide range of domains and contexts to better understand when event controllability does influence counterfactual thought.

**Aim #3: The Influence of Causal Status and Controllability on Affect Following Counterfactual Thought**

Earlier research on upward counterfactual thinking about events posited a predictive relationship between these “if only” thoughts about better outcomes and a subsequent increase in negative affect (Kahneman & Miller, 1986). To date, the controllability of events has been identified as the candidate mechanism explaining the subsequent increase in negative affect following upward counterfactual thought (Mandel, 2003b; Markman & Miller, 2006). Despite the findings of Experiments 4–6, it is possible that controllability influences changes in affect even if it does not influence the counterfactual thinking itself per se. Furthermore, given that causal status appears to influence counterfactual thought, there may also be a link between the causal status of people’s counterfactual interventions and any resulting change in affect. Experiment 7 constituted an exploratory test of whether these relationships exist. By analyzing causal maps of people’s self-reported negative events, I found that the more that people’s counterfactual thoughts focused on more causally central events, the worse that people felt after generating their counterfactuals. However, this finding was moderated by domain, as suggested by prior work (Mandel, 2003b). There was no relationship between the causal status of people’s counterfactual thoughts and subsequent affect when people wrote about academic events, but there was when they wrote about interpersonal events. These findings align with prior work suggesting that domain can affect the generation of counterfactual thought as well as the
resulting change in affect. Considerations of domain may also provide insights into the pattern of counterfactual thought for Experiments 4–6.

**Domain Effects**

The domain of the events that people reason about may also influence counterfactual thought, above and beyond any differences about possible types of causal chains. In secondary analyses, I found that people who wrote about academic events in Experiment 7 preferred to counterfactually intervene on more causally central (i.e., root causes) nodes than did people who wrote about interpersonal events. Thus, Experiment 7 explored two domains that have been shown in past work to elicit different patterns of counterfactual thought and showed for the first time that domain moderates the influence of causal status on counterfactual thought.

I further speculate that domain may moderate how causal status influences counterfactual thought in health and illness and in everyday events. Reasoning about health and disease may be one domain where there is likely to be a consistent and robust preference to intervene on the earliest, root events in a chain. People tend to assume that treating the deepest cause of a disorder is the most effective way to cure and prevent symptoms from re-occurring (Yopchick & Kim, 2009). People may also believe that expressed symptoms may not ever emerge or occur if the underlying cause was not present in the first place. It may be for this reason that people tend not to prefer treating only the surface symptoms. Many medical frameworks seem to consist of treating the deepest cause possible rather than only addressing the surface symptoms, except in cases where the cause is unknown, or untreatable with currently available interventions.

The above notions are supported by research similar to the current work. Yopchick and Kim (Experiment 3; 2009), for example, asked people to reason about fictional medical disorders (in order to minimize idiosyncratic effects of *a priori* knowledge of existing disorders) that were
characterized as a three-step causal chain with a root cause, an intermediate cause, and negative symptom as an outcome. These chains resembled the same type of event chain as used in Experiments 4–6. People were asked where in the causal chain would be most effective to intervene in order to remove the negative, outcome symptom. The authors found that even though the intermediate cause was closer in proximity to the symptom, people showed a preference to intervene on the root cause to treat the symptom. These results are consistent with the notion that the root causes are ultimately what give rise to all subsequent symptoms of a disorder. Thus, I speculate that interventional and counterfactual judgments might be more likely to consistently focus on root causes in the domain of medicine.

In contrast, reasoning about everyday events may produce different patterns of counterfactual thought compared to medical events. Everyday events can be thought of as low-stakes occurrences that may or may not happen again in the future. Research suggests that when presented with non-causal chain everyday-type occurrences and asked what would have been sufficient to prevent the outcome from occurring, people generate counterfactuals that “undo” the events immediately preceding the negative outcome (Segura & McCloy, 2003). Thus, people’s counterfactual thoughts about ordinary, everyday occurrences are likely to be concerned with undoing events in a manner that would be just sufficient to prevent the outcome that one time. (In contrast, in many diseases, a root cause such as bacteria or a tumor may be expected to continue to regenerate symptoms until eradicated.) More broadly, in order to plan for the future and achieve better outcomes when a similar situation occurs again, it is probably reasonable for people to think counterfactually about any factors, when present, that can bring about a negative outcome. Accordingly, for everyday occurrences expected to happen again, people may not necessarily feel that they need to reason exclusively about the most root cause in a chain of
events, and instead can reason about any factors sufficient to prevent the outcome. Additional work will be needed to determine whether or not these speculations about differing patterns of counterfactual thought across domains hold true.

**Causal Transitivity**

Another argument that could be made about the results of Experiments 4–6 is that people’s consistent preference to generate a counterfactual to “undo” the intermediate events in the chains was due to the critical role that intermediate events generally play in explaining how the root cause event brought about the outcome. That is, in the test scenarios used in Experiments 4–6, people may judge it essential to know what the intermediate event is in order to understand how the root event led to the outcome. People might not believe that the occurrence of the root event alone could bring about the outcome in the absence of the intermediate event. In other words, it could be argued that although the causal strength ratings between successive events in the scenarios used in Experiments 4–6 were not found to differ, the causal strength between the root event and the outcome may be relatively weaker. Yet research in causal transitivity judgments (i.e., the notion that if A → B and B → C then A → C) provides empirical evidence to counter this argument.

Specifically, Johnson and Ahn (2017) asked participants to read and reason about a series of three-step causal chains. After reading these chains, people were asked to rate the extent to which the intermediate event explains why the root event lead to the outcome (a mechanism judgment) and rate how essential it would be to mention the intermediate event when explaining to someone how the root event led to the outcome (a “chunking” judgment that reflects the extent to which the intermediate event may be implied or assumed to occur without explicitly mentioning it). Participants also made a transitivity judgment about whether the root event was a
cause of the outcome. They found that chunking judgments predicted transitivity judgments. That is, the less essential it was to mention the intermediate event, the more that the root event was judged to be a cause of the outcome. However, there was no correlation between mechanism judgments and chunking judgments. Thus, the intermediate event’s mechanistic role in explaining the root event/outcome relationship appears to be orthogonal to chunking judgments.

Given the above findings, the degree to which people believed the intermediate events in Experiments 4–6 were necessary to explain how the root event brought about the outcome (as in the mechanism judgment in Johnson & Ahn [2017]) may likewise be orthogonal to any notion of how the root event caused the outcome (as in their chunking judgments). Thus, in extrapolating to the current work, one might speculate that the counterfactual preference to undo the intermediate causes in causal chains may not be explained by the belief that the intermediate event was necessary to understand how the root cause event ultimately brought about the outcome. This, however, is an empirical question that should be addressed in future work.

**Experimental Constraints in Measuring the Influence of Controllability on Counterfactual Thought**

The influence of event controllability on counterfactual thought was well documented (e.g., Frosch et al., 2015; Mandel, 2003b; Markman & Miller, 2006) prior to the current work. However, recent work by Mercier et al. (2017) pointing to boundary conditions to this effect may be helpful in better understanding why I did not observe an effect of event controllability on counterfactual thought in Experiments 4 – 6. In a set of studies, Mercier et al. (2017) distinguished between counterfactual thought and prefactual thought (i.e., “if I leave home earlier next time, I will make my flight”). Participants in their studies completed a word-search
task for a chance to win prizes. Then, participants were either instructed to generate prefactual thoughts about how they could do better in the next round, or counterfactual thoughts that would have led to a better outcome on that round. The authors found that people in the prefactual condition generated more controllable counterfactuals than people in the counterfactual condition.

Accordingly, Mercier et al. (2017) argued that counterfactual thought that is not explicitly focused on re-occurring behaviors and events tends not to focus on the controllability of events. Specifically, Mercier et al. (2017) argued that with respect to the controllability of events, “counterfactuals may improve future performance only to the extent that they indicate the correct causal antecedent to the negative outcome” (p. 268). (In their work, “correct” causal antecedents referred to factors that had a causal rather than non-causal influence on outcomes.) That is, to improve outcomes in the future, generating counterfactual thoughts about controllable events may not be as necessary or as effective as generating them about what may be more directly causally responsible for unwanted negative outcomes. This view is consistent with people’s tendency to counterfactually “undo” the intermediate causal events immediately precipitating the outcome in Experiments 4–6.

The findings of Mercier et al. (2017) are difficult to reconcile with those of Frosch et al. (2015) and of the current Experiment 1, in which people showed a counterfactual (and not prefactual) preference for controllable events when generating counterfactuals for hypothetical events about the self and others. It may be the case that variability in the controllability of events was smaller or less salient in the Mercier et al. (2017) study than in Frosch et al. (2015) or the current Experiment 1. On the other hand, it may be that the scenarios I used in Experiments 4-6
were more conducive to prefactual thinking than the scenarios used by Frosch et al. (2015) or the current Experiment 1. Additional work is needed to determine whether this is the case.

*Controllability, Foreseeability, and Intentionality*

One particularly useful aspect of the functional framework of counterfactual thinking is that it provides a rationale for consolidating research about the level of controllability of events on counterfactual thought. However, there remain related factors such as foreseeability and intentionality that may be distinct from controllability and that may also play a functional role in counterfactual thought. Some previous research has examined how factors of foreseeability and intentionality influence causal inferences, and may therefore provide some clues toward understanding how these factors might also influence counterfactual thought. Lagnado and Channon (2008), for example, found that events that were characterized as intentional acts by people and acts that were meant to be construed as highly foreseeable both influenced judgments of cause and blame. However, a limitation of this work is that the stimulus materials were not pre-tested to ensure that people’s judgments of intentionality and foreseeability aligned with the authors’ intent. Additionally, it is unknown whether or not such judgments would be correlated with or orthogonal to judgments of controllability. Follow-up research may be needed to more clearly elucidate the differences between these factors.

If controllability, foreseeability, and intentionality are dissociable judgments that people make about events, then investigating them in tandem may provide an additional lens to better understand and resolve conflicts within the existing body of controllability-focused research. In Mercier et al.’s (2017) studies that found differences in the influence of controllability on counterfactual versus prefactual thoughts, external raters were asked to classify the generated thoughts as referring to uncontrollable or controllable aspects of a task. Future work should also
elicit judgments about whether counterfactual reasoning was influenced by the intentionality of acts or by foreseeable aspects of the task. Although reasoners generated fewer counterfactual thoughts for controllable aspects of the task than did the prefactual reasoners, they may well have generated more thoughts about intentional acts or foreseeable aspects of the task than the prefactual reasoners.

If such a pattern of results were to be found, one might still plausibly interpret it to be in line with the arguments of the functional framework of counterfactual thought. Presumably, there would be high functionality in being able to identify and think counterfactually about events in the world that one had intended to bring about, or that could be readily foreseen. Future research should be aimed at knowing whether the umbrella of the functional framework can be broadened to capture and understand the potentially related influences of such factors as intentionality and foreseeability.

**Counterfactual Thought as a Cause of Events or an Effect**

In the current work, I have reviewed literature on the factors, motivations, and forces that influence how people choose to counterfactually intervene on events. A related body of work examines the relationship between counterfactual thought and subsequent causal judgments. One central question in the latter line of work is whether counterfactual simulation about an event (rather than the mere observation of the events) can be – at times – necessary to make causal judgments about events. Some aspects of this literature may provide insights about how a much wider body of counterfactual thinking research can be synthesized to better understand the current research.

Some evidence to support the notion that counterfactual simulation can be necessary to inform judgments of cause comes from Gerstenberg, Peterson, Goodman, Lagnado, and
Tenenbaum (2017). People in their study watched computerized top-down videos of two billiard balls launched on a surface with one open gate on the far-left edge of the screen. One billiard ball was red and one was white and depending on the video, the balls would either collide or not collide with each other and then move on the resulting new paths. After viewing the same video twice, participants in their study were asked about whether the white ball would have gone through the gate if the red ball had not been launched, and responded on a numerical scale (a measure of their counterfactual judgments). Using eye-tracking devices, the authors found that during the second video, the more that people’s pattern of eye movements aligned with the counterfactual path of the white ball (i.e., where the white ball would have gone if the red ball was not launched), the more accurate (with respect to the underlying physics engine influencing the movements) their judgments were about whether the red ball caused the white ball to go through the gate.

Additionally, eye movements were collected from participants in other conditions who viewed the same videos but were asked to make judgments about whether the red ball prevented the white ball from going through the gate (a cause judgment) or who were asked to respond about whether the white ball did go through the gate (an observation judgment). The locations of eye movements on the screen were similar between counterfactual and cause judgments, whereas eye movements were not similar between the cause judgments and outcome judgments. The authors contended that such evidence supports the argument that counterfactual simulation is a critical process that people must engage in in order to make causal attributions, at least for visual physical events such as these. However, a more cautious interpretation is that whereas some form of counterfactual simulation may be necessary to make causal judgments, it does not
always follow that what is counterfactually simulated is treated as the best candidate cause for an event.

In Experiments 1–6, I intentionally manipulated the causal structure of events that were presented to people before they were asked to generate counterfactuals about the scenarios. Given Gerstenberg et al.’s (2017) arguments suggesting a strong link between counterfactual simulation and causal reasoning, there are many questions to be addressed concerning order effects between counterfactual thought and causal reasoning. Some clues about what these effects may be come from Gerstenberg, Goodman, Lagnado, and Tenenbaum (2014). A similar paradigm as in Gerstenberg et al. (2017) was employed in their work, except that the type of judgments that people were asked to make (causal or counterfactual) was manipulated within-subjects and presented in a counterbalanced order. The authors found that their counterfactual simulation model was more accurate (again, with respect to the underlying physics engine) in predicting causal judgments for people who first made counterfactual judgments than for people who first made causal judgments. Thus, recent exposure to counterfactual reasoning and simulation processes influenced the causal inferences people made about events, providing evidence for one such order effect between these judgments.

Additional studies could be conducted to determine whether there are some domains or types of causal reasoning that are not sensitive to having recently engaged in counterfactual reasoning. For example, people often form lay theories about their own health and wellness that are often not substantiated in evidence or clinical guidance. Individuals with high blood pressure often believe that other feelings and sensations (e.g., headaches) can be diagnostic of currently having high blood pressure, although no empirical evidence has validated such cues (Meyer, Leventhal, & Gotmann, 1985). Counterfactual thinking about those symptoms (e.g., “would I be
having a headache if only I had gotten more sleep last night?”) may not be sufficient to sway people’s causal inferences about the diagnosticity of their symptoms. Future research should be aimed at identifying boundary conditions for the influence of counterfactual simulation on causal inferences.

**Individual Differences**

As I discussed in the introduction, current depressive symptomology can be seen as one form of person-level characteristics that influence counterfactual thought. Individuals experiencing higher levels of depressive symptoms tend to generate a higher ratio of uncontrollable to controllable counterfactuals when reasoning about events in their own life than do those with milder or no depressive symptoms (Markman & Miller, 2006). Additionally, a meta-analysis revealed that higher depressive symptomology was associated with increased rumination about the differences between actual negative outcomes and an idealized upward counterfactual outcome (Broomhall et al., 2017). Thus, for certain individuals, increased counterfactual thought may indeed be maladaptive with respect to perpetuating their negative affect.

Individual differences in counterfactual thought may also be brought about by non-clinical factors. The ability to hold a mental representation of the world in order to manipulate and reason about it while holding it as distinct from reality – a process called **cognitive decoupling** – is proposed to be strongly linked to hypothetical thinking (Evans & Stanovich, 2013; Stanovich & Toplak, 2012). Cognitive decoupling is evident in counterfactual thinking as people form a secondary representation of reality and then mentally “undo” events and simulate how the new outcome would contrast to what actually occurred. Since counterfactual thought is a form of hypothetical thinking, the same factors that underlie individual differences in cognitive
decoupling may also produce differences between people with respect to how they reason counterfactually. Working memory capacity and fluid cognitive intelligence are both found to vary as an individual difference (Blair, 2006) and positively predict cognitive decoupling abilities (Stanovich, 2006).

As Evans and Stanovich (2013) and Stanovich and Toplak (2012) have argued that cognitive decoupling is effortful and not automatic, people who are not as likely or who would not have as great a capacity to engage in cognitive decoupling may reason counterfactually in a more automatic, low-effort manner. Whether or not counterfactual thinking can actually occur in the absence of cognitive decoupling (or in a cognitive environment bereft of the time or mental resources to meaningfully engage in it) is an empirical question that future research can address. However, the literature does provide a way to understand what low-effort counterfactual thinking may look like.

Individuals who would be generally less inclined to successfully engage in cognitive decoupling may resort to identifying elements that are more easily accessible and strongly automatically associated. Events that are low frequency, violations of norms, or not congruent with a possible schema of the focal scenario are often an automatic focus of people’s counterfactual thoughts (Kahneman & Miller, 1986). My argument is not that such factors, if influential in directing the focus of people’s counterfactual thoughts, would necessarily mislead them from identifying optimal ways to bring about better outcomes in the future. Rather, I suggest that such a pattern of thinking will not produce optimal outcomes if the best preventative strategies do not share those qualities. Future work should be aimed at providing empirical evidence about the relationship between person-level abilities in working memory capacity or fluid intelligence and how they engage in counterfactual thought.
Conclusions

Engaging in counterfactual thought is an unavoidable aspect of life for many people. What we mentally “undo” has implications for how we might direct and regulate our future behavior as well as how we make sense of our world. This latter process may be difficult and frustrating on occasions, as there is rarely an answer key for why things in the world came about, or what their true cause may be. Thus, often we may agonize over our actions or inactions in the past, and our identification of counterfactual “fault lines” that we identify in the world may intensify these feelings and increase with rumination. It is also rare that people receive feedback about whether their counterfactual thoughts are correct or not, and time is often the only way for this information to be borne out (e.g., if the situation occurs again). For these reasons, research on counterfactual thought has direct relevance to people’s lives by providing insight about what our reasoning processes may be achieving as well as what they may not be achieving toward people’s well-being. Existing research has been fruitful in this respect. We can gain insight about whether, and when, our counterfactuals reflect prevention strategies or seek to identify causes (Mandel, 2003b). We can also become better aware of how we go about assigning blame or responsibility to others (Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Zultan, Gerstenberg, & Lagnado, 2012).

The current work provides evidence that when reasoning about everyday ordinary events, people have a tendency to counterfactually intervene on intermediate causes in proximity to a negative outcome. As of now, it is unknown how this general tendency predicts whether successful or unsuccessful outcomes are achieved in the future. Future work is needed to link this pattern of counterfactual thought to future outcomes, and to contrast it with the actual underlying causes of events. People are likely quite aware of what their counterfactuals focus
on. My speculation is that the world does not operate in such a way that people’s counterfactual thoughts about the world are likely to be consistently constrained or manipulated to ensure the best outcomes later in life. A worthy goal of cognition research, then, is to help people become aware of the limitations as well as the utility that their counterfactual thoughts afford.
References


Gelman, A. (2018, March 15). You need 16 times the sample size to estimate an interaction than to estimate a main effect. Retrieved from:


http://dx.doi.org/10.1037/xge0000556


Footnotes

1 Kai Epstude and Neal Roese have together written comprehensive reviews on counterfactual thinking’s relationship to goal cognition and future successful behavior under the umbrella of what they refer to as the functional theory of counterfactual thinking (see Epstude & Roese, 2008; Roese & Epstude, 2017). However, multiple other researchers have summarized the state of self-relevant counterfactual thinking research without claiming the findings speak for or against a specific theory. My choice to refer to the functional framework of counterfactual thinking aligns more with other researchers who have similarly summarized the literature as describing a functional side or perspective of counterfactual thinking (Markman, Karadogan, Lindberg, & Zell, 2008) or who have synthesized prior work without positing that any one specific theory accounts (or is needed to account) for how counterfactual thinking is linked to planning and reasoning about one’s future behavior (Byrne, 2016). Accordingly, in this paper I described the functional perspective as a loose framework for organizing and understanding prior research. The current work is not concerned with providing evidence for or against any singular theory of counterfactual thinking (cf., Epstude & Roese, 2008; Roese & Epstude, 2017).

2 The intentionality and the foreseeability of the events were also manipulated across several of their experiments. In the General Discussion, I speculate on whether conceptual similarities as well as differences between intentionality and foreseeability and controllability might be shown to impact the current results in future work.

3 Including these participants in the analyses did not change the results.

4 It could be interpreted that our manipulation of the first-person perspective was not strong enough to evoke real first-person counterfactual reasoning, as participants in this condition were reasoning about hypothetical scenarios provided to them by the experimenter.
Although a more individualized design to elicit a first-person perspective in participants should be explored in future studies, the benefit of the current designs is that the same information is provided to participants in the first-person as in the third-person perspective. Manipulating the level of involvement in a scenario by having participants act out rather than read parts makes different information salient to the participant (Giorotto, Ferrante, Pighin, & Gonzales, 2007) and thus prevents information from being consistent across perspectives. In Experiment 7, I took a first step towards understanding the relationship between event controllability (as well as causal status) and counterfactual thought when people reason about negative events that have occurred to them in their own lives.

A limitation of this design is that people only reason about one general event and the extent to which the study findings generalize to more situations or domains is less clear. However, this design has the advantage of avoiding confounding causal position with content (as each event will appear equally as often in each causal position) while maintaining the overall coherence of the story.

It also must be noted that the coding process between the independent coders and the participants themselves was necessarily different. The independent coders were able to see a variety of counterfactual responses across the studies. The participants themselves only encountered their own counterfactuals, and were thus not able to calibrate their coding decisions based on the responses and coding decisions of others.

It is the relative difference between the causal weights in any given causal map that affects the computed causal centrality of the features. Thus assigning all causal links in a given map a weight of 1 should produce no difference in the resulting order of causal centrality scores
compared to assigning all links a weight of 3. The resulting causal centrality scores for nodes (event features) are only meaningful on a rank order scale and not ratio scale.
Figure 1. Sample causal chain structure. Event A tends to cause event B, which in turn tends to cause event C, which in turn tends to cause event D, resulting in outcome X.
Figure 2. Sample conjunctive common effect structure. Events A, B, C, and D unfold consecutively over time without any causal relationships directly linking them to each other. The occurrence of all four of these events causes outcome X to occur. The occurrence of only one, two, or three events would not be enough to cause outcome X; only the occurrence of all four events together would result in outcome X.
Figure 3a. Sample visual diagram of one of the causal chain scenarios used in Experiment 2.
Figure 3b. Sample visual diagram of one of the conjunctive common-effect scenarios used in Experiment 2.
Figure 4. Causal strength ratings by causal link and intended controllability for the six-scenario balanced subset identified in Experiment 3b. On the y-axis, the larger the number, the more strongly people rated the first event in the causal pair as a cause of the second event.
Figure 5. Controllability ratings by causal status and intended controllability for the six-scenario balanced subset identified in Experiment 3b. On the y-axis, the larger the number, the more the event was judged to be controllable by the focal actor.
Figure 6. The focus of people’s counterfactual thoughts by event controllability and causal status in Experiment 4. Bars reflect the percentage of scenarios for each event type (e.g., controllable root cause and uncontrollable intermediate cause) to which participants directed the focus of their counterfactual thoughts. Error bars reflect standard errors.
Figure 7. The focus of people’s counterfactual thoughts by event controllability and causal status collapsed across counterfactual prompt in Experiment 5. Bars reflect the percentage of scenarios for each event type (e.g., controllable root cause and uncontrollable intermediate cause) to which participants directed the focus of their counterfactual thoughts. Error bars reflect standard errors.
Figure 8. The focus of people’s counterfactual thoughts by event controllability and causal status in Experiment 6. Bars reflect the percentage of scenarios for each event type (e.g., controllable root cause and uncontrollable intermediate cause) to which participants directed the focus of their counterfactual thoughts. Error bars reflect standard errors.
<table>
<thead>
<tr>
<th>Effect</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>$\eta^2_p$</th>
<th>BF$^{a}_{inclusion}$</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point of view</td>
<td>1, 126</td>
<td>.62</td>
<td>.434</td>
<td>.01</td>
<td>.054</td>
<td>Strong H$^0$</td>
</tr>
<tr>
<td>Controllability</td>
<td>1, 126</td>
<td>62.61</td>
<td>&lt; .001</td>
<td>.33</td>
<td>4.408e + 6</td>
<td>Extreme H$_1$</td>
</tr>
<tr>
<td>Point of view × Controllability</td>
<td>1, 126</td>
<td>.26</td>
<td>.612</td>
<td>&lt; .01</td>
<td>.035</td>
<td>Strong H$^0$</td>
</tr>
<tr>
<td>Event Number</td>
<td>1, 126</td>
<td>3.51</td>
<td>.064</td>
<td>.03</td>
<td>.255</td>
<td>Moderate H$^0$</td>
</tr>
<tr>
<td>Point of view × Event Number</td>
<td>1, 126</td>
<td>1.21</td>
<td>.274</td>
<td>.01</td>
<td>.027</td>
<td>Very Strong H$^0$</td>
</tr>
<tr>
<td>Controllability × Event Number</td>
<td>1, 126</td>
<td>1.87</td>
<td>.174</td>
<td>.02</td>
<td>.425</td>
<td>Anecdotal</td>
</tr>
<tr>
<td>Point of view × Controllability × Event Number</td>
<td>1, 126</td>
<td>1.51</td>
<td>.222</td>
<td>.01</td>
<td>.013</td>
<td>Very Strong H$^0$</td>
</tr>
</tbody>
</table>

$^a$ BF$^{inclusion}$ reflects support for including that effect in the model
Table 2

Frequency of counterfactual choice by event position and causal structure in Experiment 2

<table>
<thead>
<tr>
<th>Event choice</th>
<th>First Event</th>
<th>Second Event</th>
<th>Third Event</th>
<th>Fourth Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causal Chain</td>
<td>39</td>
<td>22</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Conjunctive Common Effect</td>
<td>21</td>
<td>16</td>
<td>15</td>
<td>19</td>
</tr>
</tbody>
</table>
Table 3
Percentage of Experiment 4 participants generating counterfactuals by node and scenario

<table>
<thead>
<tr>
<th>Node that was “undone” in counterfactual thought</th>
<th>Scenario type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uncontrollable root cause → controllable intermediate cause → outcome</td>
</tr>
<tr>
<td></td>
<td>Controllable root cause → uncontrollable intermediate cause → outcome</td>
</tr>
<tr>
<td></td>
<td>Scenario number</td>
</tr>
<tr>
<td></td>
<td>#1</td>
</tr>
<tr>
<td>Root cause only</td>
<td>12.5</td>
</tr>
<tr>
<td>Intermediate cause only</td>
<td>37.5</td>
</tr>
<tr>
<td>Both root and intermediate cause</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>50.0</td>
</tr>
</tbody>
</table>
Table 4. Analysis of Variance Results in Experiment 4 for Causal Status and Controllability

<table>
<thead>
<tr>
<th>Effect</th>
<th>$df$</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_p$</th>
<th>$BF^a_{\text{inclusion}}$</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causal Status</td>
<td>1, 79</td>
<td>94.41</td>
<td>&lt; .001</td>
<td>.54</td>
<td>6.005e + 15</td>
<td>Extreme $H_1$</td>
</tr>
<tr>
<td>Controllability</td>
<td>1, 79</td>
<td>.27</td>
<td>.607</td>
<td>&lt; .01</td>
<td>.117</td>
<td>Moderate $H_0$</td>
</tr>
<tr>
<td>Causal Status $\times$ Controllability</td>
<td>1, 79</td>
<td>1.18</td>
<td>.280</td>
<td>.02</td>
<td>.120</td>
<td>Moderate $H_0$</td>
</tr>
</tbody>
</table>

$^a$ $BF_{\text{inclusion}}$ reflects support for including that effect in the model
Table 5

Percentage of Experiment 5 participants in the sentence-completion prompt condition generating counterfactuals by node and scenario

<table>
<thead>
<tr>
<th>Node that was “undone” in first counterfactual thought</th>
<th>Scenario type</th>
<th>Scenario number</th>
<th>Scenario number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uncontrollable root cause → controllable intermediate cause → outcome</td>
<td></td>
<td>Controllable root cause → uncontrollable intermediate cause → outcome</td>
</tr>
<tr>
<td></td>
<td>Scenario number</td>
<td></td>
<td>Scenario number</td>
</tr>
<tr>
<td></td>
<td>#1</td>
<td>#2</td>
<td>#3</td>
</tr>
<tr>
<td>Root cause only</td>
<td>24.7</td>
<td>18.5</td>
<td>14.8</td>
</tr>
<tr>
<td>Intermediate cause only</td>
<td>27.2</td>
<td>24.7</td>
<td>49.4</td>
</tr>
<tr>
<td>Both root and intermediate cause only</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>48.1</td>
<td>56.8</td>
<td>35.8</td>
</tr>
</tbody>
</table>
Table 6

Percentage of Experiment 5 participants in the open-ended prompt condition generating counterfactuals by node and scenario

<table>
<thead>
<tr>
<th>Node that was “undone” in first counterfactual thought</th>
<th>Scenario type</th>
<th>Uncontrollable root cause → controllable intermediate cause → outcome</th>
<th>Controllable root cause → uncontrollable intermediate cause → outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root cause only</td>
<td>#1</td>
<td>21.5</td>
<td>7.6</td>
</tr>
<tr>
<td>Intermediate cause only</td>
<td>#2</td>
<td>10.1</td>
<td>21.5</td>
</tr>
<tr>
<td>Both root and intermediate cause</td>
<td>#3</td>
<td>10.1</td>
<td>15.2</td>
</tr>
<tr>
<td>Other</td>
<td>#4</td>
<td>38.0</td>
<td>62.0</td>
</tr>
<tr>
<td></td>
<td>#5</td>
<td>49.4</td>
<td>26.6</td>
</tr>
<tr>
<td></td>
<td>#6</td>
<td>60.8</td>
<td>34.2</td>
</tr>
<tr>
<td></td>
<td>#4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>#5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>#6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 7. Analysis of Variance Results in Experiment 5 for Causal Status, Controllability, and Counterfactual Prompt

<table>
<thead>
<tr>
<th>Effect</th>
<th>df</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2$</th>
<th>BF$_{\text{inclusion}}$</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causal Status</td>
<td>1, 158</td>
<td>86.69</td>
<td>&lt; .001</td>
<td>.34</td>
<td>$\infty$</td>
<td>Extreme H$_1$</td>
</tr>
<tr>
<td>Controllability</td>
<td>1, 158</td>
<td>.08</td>
<td>.784</td>
<td>&lt; .01</td>
<td>.073</td>
<td>Strong H$_0$</td>
</tr>
<tr>
<td>Causal Status × Controllability</td>
<td>1, 158</td>
<td>1.19</td>
<td>.28</td>
<td>.01</td>
<td>.056</td>
<td>Strong H$_0$</td>
</tr>
<tr>
<td>Prompt</td>
<td>1, 158</td>
<td>1.72</td>
<td>.191</td>
<td>.01</td>
<td>.371</td>
<td>Anecdotal</td>
</tr>
<tr>
<td>Causal Status × Prompt</td>
<td>1, 158</td>
<td>3.78</td>
<td>.054</td>
<td>.02</td>
<td>1.412</td>
<td>Anecdotal</td>
</tr>
<tr>
<td>Controllability × Prompt</td>
<td>1, 158</td>
<td>6.51</td>
<td>.012</td>
<td>.04</td>
<td>.219</td>
<td>Moderate H$_0$</td>
</tr>
<tr>
<td>Causal Status × Controllability × Prompt</td>
<td>1, 158</td>
<td>2.86</td>
<td>.093</td>
<td>.02</td>
<td>.050</td>
<td>Strong H$_0$</td>
</tr>
</tbody>
</table>

$^a$ BF$_{\text{inclusion}}$ reflects support for including that effect in the model.
Table 8

Percentage of Experiment 6 participants coding their own counterfactuals by node and scenario

<table>
<thead>
<tr>
<th>Node that was “undone” in counterfactual thought</th>
<th>Scenario type</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Uncontrollable root cause → controllable intermediate cause → outcome</td>
<td>#1</td>
<td>#2</td>
<td>#3</td>
<td>#4</td>
<td>#5</td>
<td>#6</td>
</tr>
<tr>
<td>Root cause only</td>
<td></td>
<td>19.5</td>
<td>12.2</td>
<td>15.9</td>
<td>14.6</td>
<td>24.4</td>
<td>15.9</td>
</tr>
<tr>
<td>Intermediate cause only</td>
<td></td>
<td>65.9</td>
<td>47.6</td>
<td>73.2</td>
<td>74.4</td>
<td>56.1</td>
<td>47.6</td>
</tr>
<tr>
<td>Both root and intermediate cause</td>
<td></td>
<td>7.3</td>
<td>6.1</td>
<td>8.5</td>
<td>4.9</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>7.3</td>
<td>34.1</td>
<td>2.4</td>
<td>6.1</td>
<td>17.1</td>
<td>34.1</td>
</tr>
</tbody>
</table>
Table 9. Analysis of Variance Results in Experiment 6 for Causal Status and Controllability

<table>
<thead>
<tr>
<th>Effect</th>
<th>df</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2$</th>
<th>BF_{inclusion}</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causal Status</td>
<td>1, 81</td>
<td>91.45</td>
<td>&lt; .001</td>
<td>.53</td>
<td>$\infty$</td>
<td>Extreme $H_1$</td>
</tr>
<tr>
<td>Controllability</td>
<td>1, 81</td>
<td>.56</td>
<td>.456</td>
<td>&lt; .01</td>
<td>.140</td>
<td>Moderate $H_0$</td>
</tr>
<tr>
<td>Causal Status $\times$ Controllability</td>
<td>1, 81</td>
<td>.02</td>
<td>.901</td>
<td>&lt; .01</td>
<td>.102</td>
<td>Moderate $H_0$</td>
</tr>
</tbody>
</table>

$^a$ BF_{inclusion} reflects support for including that effect in the model
Table 10. Mean affect ratings by time and negative event context in Experiment 7.

<table>
<thead>
<tr>
<th>Event context</th>
<th>Pre-counterfactual affect rating</th>
<th>Post-counterfactual affect rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Academic event</td>
<td>16.80</td>
<td>6.17</td>
</tr>
<tr>
<td>Interpersonal event</td>
<td>17.33</td>
<td>5.65</td>
</tr>
</tbody>
</table>

Note: Possible range of affect ratings was from 3 to 27 with higher values indicating more negative affect.
Appendix A

Procedure Used in Experiment 3b to Identify Balanced Stimulus Materials

The 16 scenarios rated for perceived controllability and causal strength in Experiment 3b could not be analyzed with scenario as a within-subject factor, as scenario always corresponded to a specific arrangement of intended controllability and causal link. Thus, ratings were collapsed across scenarios of the same arrangement. Subsets of scenarios were tested by averaging a subset of scenarios from each type (i.e., from controllable root cause/uncontrollable intermediate cause scenarios and from uncontrollable root cause/controllable intermediate cause scenarios).

I found that no seven, six, five, or four-event scenario subsets (picked from each chain type) could be balanced for causal strength ratings and intended controllability. There was a successful three-event subset (i.e., six events total) that was balanced for causal strength ratings and intended controllability. Those results are reported in Experiment 3b and the scenarios are printed verbatim in Appendix B.
Appendix B

Scenarios Used in Experiments 4–6

This appendix contains the scenarios used in the main task in Experiments 4 through 6. The scenarios are organized by whether they contain an uncontrollable event causing a controllable event causing a negative outcome or whether they contain a controllable event causing an uncontrollable event causing a negative outcome. Within each scenario, events meant to be uncontrollable are italicized and the events meant to be controllable are in boldface.

Uncontrollable Root Cause and Controllable Intermediate Cause

1. I was completing errands one day when the following events occurred: I was in my car driving down the road. I swerved to avoid hitting a dog that suddenly jumped into the road, and I ended up on the curb. I decided to back up rather than drive forward to get back on the road. This, in turn, caused me to get into a minor accident with a parked car.

2. I was travelling for my yearly vacation when the following events occurred: A flight back home from vacation was delayed due to bad weather. Because of this, I decided to walk around the airport to keep busy. Because of this, I slipped on a puddle of water while walking around and sprained my ankle.

3. I was returning from shopping one day when the following events occurred: A sign and a barrier forced me to take the scenic route home rather than my usual route. This led me to drive through a yellow light to make up for lost time, which led me to get into an accident with another car.

Controllable Root Cause and Uncontrollable Intermediate Cause
4. After I decided on my plans for the weekend the following events occurred: I visited a brand new restaurant in my neighborhood. *Because the restaurant was new and unfamiliar to me, the waitstaff insisted on bringing me the house special.* I did not like the dish and was disappointed with my experience.

5. My old cell phone contract expired and the following events occurred: I decided to change cell phone plans in order to save money. *This caused me to visit a store where the newly hired service rep inadvertently changed my phone number.* This caused me to miss a text about a friend’s surprise birthday party while the rep was still trying to change it back.

6. My apartment lease was about to expire and the following events occurred: I decided to move to a different apartment so that I would have more space. *To do so, I had to hire a moving crew, as this was the policy of the new complex.* One of the workers from the moving company either took or lost a box of expensive electronics during the move.
Appendix C

Combined Analyses of Participant-Coded and Independent Rater-Coded Counterfactuals

In Experiment 6

Participants were asked to code their own counterfactual statements in Experiment 6 to see whether the results of Experiment 4 would replicate when people (rather than external others) classified their own counterfactual responses. As seen in Experiment 6, the results replicated the preference to “undo” intermediate over root causes, and there was no main effect of event controllability or interaction of event controllability and causal status.

I also carried out a secondary analysis of people’s counterfactual responses in Experiment 6. As in Experiments 4 and 5, I asked external coders to likewise classify the participant’s counterfactuals to determine more directly whether coding source could have moderated these findings.

The same two coders from Experiments 4 and 5 coded the participants’ counterfactual responses from Experiment 6. At the time that they carried out the coding, the coders were not made aware of the results of the participants’ own coding. The coders agreed on 89% of all counterfactuals and resolved all discrepancies by discussion. As in Experiments 4 and 5, for each participant, I calculated the percentage of scenarios in which coders classified the counterfactual as focusing primarily on the controllable-root, uncontrollable-root, controllable-intermediate, and uncontrollable intermediate causes.

To examine whether any of our prior findings were moderated by the source of the coding, I analyzed the results from the participants’ own coding along with the independent coders’ results with source of coding as a within-subject factor. The codes made by the independent raters aligned with 65.04% the codes of the participants.
I conducted a 2 (causal status: root cause, intermediate cause) x 2 (controllability: uncontrollable, controllable) x 2 (coder: self, other) repeated-measures ANOVA on these percentages. Once again I found a main effect of causal status, $F(1, 81) = 73.51, p < .001, \eta_p^2 = .48$. Counterfactuals were coded as referring to the intermediate cause event in a higher percentage of scenarios ($M = 50.20\%, SD = 19.26$) than to root cause events ($M = 15.96\%, SD = 20.79$), and this pattern occurred across all six scenarios. There was no main effect of controllability, $F(1, 81) = 0.02, p = .129, \eta_p^2 = .72$. There was also no interaction of causal status and controllability, $F(1, 81) = 2.11, p = .149, \eta_p^2 = .03$.

There was a main effect of coder, $F(1, 81) = 80.78, p < .001, \eta_p^2 = .50$. Participants said their counterfactual statements referred to either the root or intermediate cause in a higher percentage of scenarios overall ($M = 38.92\%, SD = 9.32$) than did the independent coders ($M = 27.24\%, SD = 11.86$). This pattern of results may suggest that the independent coders were conservative in making their classifications, only classifying statements that clearly described undoing either the root or intermediate cause as such.\textsuperscript{6}

There was an interaction of causal status and coder, $F(1, 81) = 33.70, p < .001, \eta_p^2 = .29$. The two coding sources did not differ in the percentage of scenarios coded that involved the root cause ($t(81) = 1.62, p = .109, d = .18$), but participants themselves said they generated counterfactuals about intermediate events in a higher percentage of scenarios ($M = 60.77\%, SD = 24.19$) than did the independent coders ($M = 39.63\%, SD = 20.76$), $t(81) = 8.12, p < .001, d = .90$.

There was also an interaction of controllability and coder, $F(1, 81) = 9.00, p = .004, \eta_p^2 = .10$. There was no preference to undo controllable compared to uncontrollable events for either participants ($t(81) = 0.75, p = .456, d = .08$) or independent coders ($t(81) = 1.54, p = .128, d = $
However, participants classified a higher percentage of their counterfactuals as focusing on uncontrollable events ($M = 37.60\%, SD = 19.74$) than did the independent coders ($M = 29.67\%, SD = 18.52$), $t(81) = 4.03, p < .001, d = 1.38$. Participants also classified a higher percentage of their counterfactuals as focusing on controllable events ($M = 40.24\%, SD = 17.16$) than did the independent coders ($M = 24.80\%, SD = 18.75$), $t(81) = 9.47, p < .001, d = 1.78$. Thus, these latter two results appear to reflect a generally conservative coding process conducted by the independent coders.

Finally, there was a three-way interaction of controllability, causal status, and coder, $F(1, 81) = 4.81, p = .031, \eta^2_p = .056$. To more clearly unpack the nature of these interactions, I also analyzed the results broken down by coder. That is, I separately examined the participants’ self-coding results and the independent coders’ results. As described in the analyses below, this three-way interaction appears to reflect an interaction of causal status and controllability that was present in the independently coded counterfactuals and absent in the participant-coded counterfactuals.

**Participants’ self-coding results.** A 2 (causal status: root cause, intermediate cause) x 2 (controllability: uncontrollable, controllable) repeated-measures ANOVA was conducted on the same counterfactual percentages. The results revealed, once again, a main effect of causal status, $F(1, 81) = 91.45, p < .001, \eta^2_p = .53$. People classified a larger percentage of their own counterfactual statements as referring to intermediate causes ($M = 60.77\%, SD = 24.45$) than root causes ($M = 17.02\%, SD = 20.83$). Once again, this pattern was consistent across all six scenarios about which people reasoned. There was no effect of controllability, $F(1, 81) = .56, p = .46, \eta^2_p = .01$ and no interaction of causal status and controllability, $F(1, 81) = .02, p = .901, \eta^2_p < .01$. 

143
**Independent coders’ results.** A 2 (causal status: root cause, intermediate cause) x 2 (controllability: uncontrollable, controllable) repeated-measures ANOVA was conducted on the above percentages. The results revealed, once again, a main effect of causal status, $F(1, 81) = 40.16, p < .001$, $\eta^2_p = .33$. Counterfactuals referred to intermediate causes in a higher percentage of scenarios ($M = 40.12, SD = 20.54$) than to root causes ($M = 14.61, SD = 22.72$); the direction of means was consistent across all six scenarios. There was no main effect of controllability, $F(1, 81) = 2.36, p = .128$, $\eta^2_p = .03$. However, there was, for the first time, an interaction of controllability and causal status, $F(1, 81) = 5.08, p = .027$, $\eta^2_p = .06$.

I ran four Bonferroni-corrected paired comparisons ($\alpha = .0125$) to more closely examine the interaction. People generated counterfactuals about uncontrollable events in a higher percentage of scenarios when they occupied an intermediate ($M = 44.31\%, SD = 29.19$) compared to root causal status ($M = 15.04\%, SD = 26.79$), $t(81) = 6.31, p < .001, d = .70$. Likewise, people generated counterfactuals about controllable events in a larger percentage of scenarios when they were intermediate ($M = 34.96\%, SD = 26.24$) compared to root causes ($M = 14.63\%, SD = 26.24$), $t(81) = 4.66, p < .001, d = .51$. However, there was no difference between the percentage of scenarios for which people generated counterfactuals for root causes when they were controllable ($M = 14.63\%, SD = 26.24$) versus uncontrollable ($M = 15.04\%, SD = 26.79$), $t(81) = .13, p = .894, d = .02$. There was also no difference between intermediate causes when they were controllable ($M = 34.96\%, SD = 28.16$) versus uncontrollable ($M = 44.31\%, SD = 29.19$), $t(81) = 2.16, p = .034, d = .07$.

**Summary.** The results from this secondary, combined analysis suggest that the independent coders were generally more conservative in their coding compared to the people coding their own counterfactuals. However, such a pattern does suggest that the participants
were well-able to complete the task and meaningfully classify their counterfactual responses. Although I did find, for the first time, an interaction of causal status and controllability for the independent coder data, the Bonferroni-corrected pairwise comparisons only reflected the counterfactual preference for intermediate over root events consistently found in Experiments 4 and 5. Taken together, therefore, the data collected throughout the current work do not reliably suggest that causal status and controllability interactively influence counterfactual thought.