COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2016) 1–13

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12451

Counterfactual Plausibility and Comparative Similarity

Matthew L. Stanley,[a] Gregory W. Stewart,[b,c] Felipe De Brigard[a,b,c]

[a]*Center for Cognitive Neuroscience, Duke University*
[b]*Department of Philosophy, Duke University*
[c]*Duke Institute for Brain Sciences, Duke University*

Received 26 April 2016; received in revised form 5 August 2016; accepted 31 August 2016

Abstract

Counterfactual thinking involves imagining hypothetical alternatives to reality. Philosopher David Lewis (1973, 1979) argued that people estimate the subjective plausibility that a counterfactual event might have occurred by comparing an imagined possible world in which the counterfactual statement is true against the current, actual world in which the counterfactual statement is false. Accordingly, counterfactuals considered to be true in possible worlds comparatively more similar to ours are judged as more plausible than counterfactuals deemed true in possible worlds comparatively less similar. Although Lewis did not originally develop his notion of comparative similarity to be investigated as a psychological construct, this study builds upon his idea to empirically investigate comparative similarity as a possible psychological strategy for evaluating the perceived plausibility of counterfactual events. More specifically, we evaluate judgments of comparative similarity between episodic memories and episodic counterfactual events as a factor influencing people's judgments of plausibility in counterfactual simulations, and we also compare it against other factors thought to influence judgments of counterfactual plausibility, such as ease of simulation and prior simulation. Our results suggest that the greater the perceived similarity between the original memory and the episodic counterfactual event, the greater the perceived plausibility that the counterfactual event might have occurred. While similarity between actual and counterfactual events, ease of imagining, and prior simulation of the counterfactual event were all significantly related to counterfactual plausibility, comparative similarity best captured the variance in ratings of counterfactual plausibility. Implications for existing theories on the determinants of counterfactual plausibility are discussed.

Keywords: Counterfactual; Plausibility; Simulation; Episodic; Autobiographical; Similarity

Correspondence should be sent to Felipe De Brigard, Department of Philosophy, Duke University, 203A West Duke Building, Durham, NC 27708. E-mail: felipe.debrigard@duke.edu

1. Introduction

Our tendency to revisit events from the personal past and to imagine alternative ways in which they might have occurred instead—a class of hypothetical thinking known as *episodic counterfactual thought* (De Brigard & Giovanello, 2012)—is both pervasive and useful in everyday life (Byrne, 2005, 2016; Epstude & Roese, 2008). Episodic counterfactual thoughts typically depict either better ("upward counterfactuals") or worse ("downward counterfactuals") ways in which past events might have occurred (Epstude & Roese, 2008). Despite the fact that people know that the events simulated during episodic counterfactual thinking did not occur, they nonetheless frequently judge some imagined counterfactual events as more plausible than others. However, the reasons as to why certain imagined counterfactual events are considered more or less plausible remain poorly understood.

In his seminal work on counterfactuals, philosopher David Lewis (1973) argued that, in evaluating the plausibility of a counterfactual statement, we compare a possible world in which the counterfactual statement is true against the current, actual world in which the counterfactual statement does not obtain. Counterfactuals considered to be true in possible worlds comparatively closely similar to ours are judged as more plausible than counterfactuals deemed true in possible worlds comparatively less similar. Philosophers and logicians have largely interpreted Lewis's notion of comparative similarity as a way to determine the semantic value of counterfactual statements (Bowie, 1979; Fine, 1975; Williamson, 1988). Despite it being interpreted as such, and regardless of Lewis's original intent, a version of Lewis's notion of comparative similarity can be usefully studied as a psychological construct. More specifically, individuals may, as a matter of psychological fact, make a context-sensitive comparative similarity judgment between actual and counterfactual events to evaluate the perceived plausibility of counterfactual events having occurred. Unfortunately, this psychological interpretation of comparative similarity as a strategy for evaluating the perceived plausibility of counterfactual events has not been empirically scrutinized. Instead, most empirical research on counterfactual thinking has focused on other factors that may influence the perceived plausibility of imagined counterfactual events.

Seminal psychological research on biases and heuristics led to one of the first and most influential views on the determinants of counterfactual plausibility (Kahneman & Miller, 1986; Kahneman & Tversky, 1982). The simulation heuristic refers to the tendency for people to determine the subjective plausibility of a hypothetical scenario occurring based on the ease with which the hypothetical event is imagined. Kahneman and Tversky (1982) presented participants with a scenario describing two individuals who arrive at the airport expecting to miss their flights. Passenger A missed his scheduled flight because he was 30 min late to the airport. However, passenger B missed his flight only by 5 min because his flight had been delayed. Even though both men missed their flights, participants thought that passenger B would experience more regret for having missed his flight relative to passenger A. This is the case—the authors suggest—because participants judge that it would have been more plausible for passenger B to make his flight than for passenger A, since it is easier to construct a mental simulation in which passenger B makes the flight.

Inspired by prior work demonstrating that the repeated simulation of possible future events tends to affect judgments of their perceived plausibility (Koehler, 1991; Szpunar & Schacter, 2013), recent empirical evidence has shown that the repeated simulation of counterfactual events reduces their perceived novelty, which in turn renders the imagined counterfactual event as seemingly less plausible (De Brigard, Szpunar, & Schacter 2013). These prior results support the mental models view on counterfactual generation (Byrne, 1997, 2002; Johnson-Laird & Byrne, 2002). According to this view, when people naturally generate counterfactual thoughts, they tend to contrast a template mental representation of what is thought to be true or normal against another mental representation that is thought to be false but deviates minimally from the template. Accordingly, the less the counterfactual alternative deviates from the template, the more plausible it seems. In the case of episodic counterfactual thoughts, if the initially generated counterfactual thought diverges minimally from the actual memory, the perceived plausibility of the counterfactual event should be high. But when repeated simulation allows for more attention to be allocated to survey discrepancies between the actual and the counterfactual mental model, the counterfactual event starts to seem less plausible.

The purpose of this study is to evaluate comparative similarity as a factor influencing people's judgments of plausibility in counterfactual simulations, and also to compare it against other factors that may influence judgments of plausibility. We limit our investigation to episodic counterfactual thoughts, as the memory of the actual event provides a reasonable standard against which to contrast counterfactual events. We investigate five main questions. First, we investigate whether greater perceived similarity between the original memory and the counterfactual event is, in fact, associated with greater perceived plausibility of the counterfactual event. Second, we investigate whether novel counterfactual simulations (i.e., *not* previously simulated outside of the laboratory setting) seem more plausible. Third, we investigate whether increased ease of imagining is associated with increased perceived plausibility. Fourth, we investigate whether the magnitude of the correlation between similarity and plausibility is significantly greater than the magnitude of the correlation between either ease of imagining and plausibility or counterfactual novelty and plausibility (or both). Fifth, using model fit procedures, we further investigate which combination of these variables (if any) together best explains the variability in perceived counterfactual plausibility.

2. Method

2.1. Participants

A total of 49 people participated in this study. Three participants were removed due to computer malfunction and two participants misunderstood instructions, so data from 44 participants were analyzed (M_{age} = 22.37, SD = 2.96; 27 females). Written informed consent was obtained from each participant in accordance with protocol approved by the Duke University Campus Institutional Review Board.

3. Materials and procedure

The experimental paradigm employed in this study (Fig. 1) was adapted from De Brigard et al. (2013) and Szpunar and Schacter (2013). This study consisted of three sessions. In session 1, participants visited the laboratory to generate 17 negative and 17 positive autobiographical memories of specific events that they had personally experienced within the past 5 years. Participants were asked to be as specific as possible and to

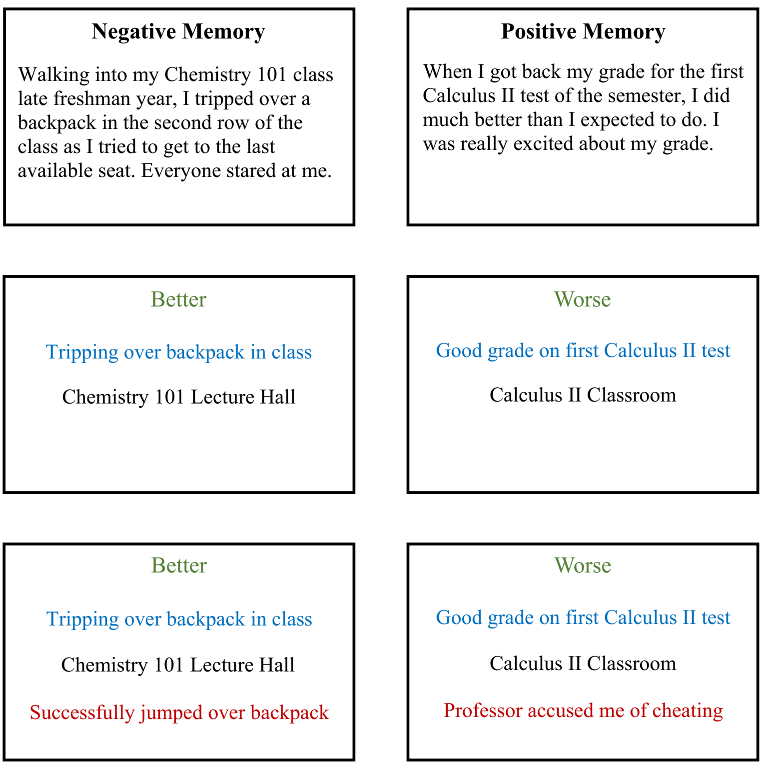

Figure 1. Examples of autobiographical memories reported in Session 1 and the corresponding stimulus displays and counterfactual generations from the following sessions. The top row shows examples of negative and positive autobiographical memories. The second row shows the stimulus displays used to prompt the generation of upward ("Better") and downward ("Worse") counterfactual simulations for these events. During Session 2, participants saw the brief title (shown in blue) of each memory on separate displays, the location where it took place, and a header on the screen (shown in green) indicating the direction of the counterfactual. Participants generated upward and downward counterfactuals for both positive and negative memories. The third row illustrates examples of counterfactuals generated by participants that were then re-presented to them during Session 3; in each case, the new title given to the counterfactual event is shown in red.

avoid events that blend into other similar events. For each memory, participants typed a short description of the memory (2–4 sentences), as well as the location where the event took place and a brief title for the memory. Participants were told that the brief title and location needed to be specific enough such that they would be able to recollect that same exact memory in session 2 if cued with those two pieces of information. Participants also provided more specific ratings of valence (1 = *very negative*, 7 = *very positive*) for each memory.

One week later, participants returned to the lab for an episodic counterfactual simulation session. Participants were told that they would be engaging in two different kinds of counterfactual simulations—better (upward) and worse (downward) counterfactual alternatives to their actual memories—and that each simulation would be prompted by a display on a computer monitor. When the heading on the screen was "Better," they would be presented with the brief title and location of a memory from the first session, and they would be asked to imagine a better way in which the cued memory might have occurred. Half of the cued memories from session 1 were negative, and the other half were positive. When the heading on the screen was "Worse," they would be presented with the brief title and location of a memory from the first session, and they would be asked to imagine a worse way in which the cued memory might have occurred. Half of the cued memories from session 1 were negative, and the other half were positive. Exactly one counterfactual was generated for each individual memory. Participants were allotted 12.5 s to simulate each episodic counterfactual event. After each imagining, they were prompted to write a short new title for the simulated counterfactual event. To ensure that participants understood all instructions, we conducted two practice trials in which participants simulated one downward and one upward episodic counterfactual event randomly chosen from the set of memories provided during session 1. Materials were presented with E-Prime software (Psychology Software Tools, Pittsburgh, PA, USA), and participants used a keyboard to type the titles for the simulated counterfactual events.

The third and final session took place 1 day later. Participants were asked to re-simulate the counterfactuals that they had generated during the second session. For each individual stimulus display, participants were given the brief title of the original memory provided during session 1, the location of the original memory provided during session 1, and the title for the counterfactual event provided during session 2. All counterfactuals were re-presented to participants in random order. Participants were given 12.5 s to re-simulate each event. Participants were explicitly asked to re-simulate the very same counterfactuals they had generated the day before. After each individual counterfactual event was re-simulated, participants provided a set of phenomenological ratings for each counterfactual, according to their ease of imagining (1 = *very difficult to imagine*, 7 = *very easy to imagine*), valence (1 = *very negative*, 7 = *very positive*), plausibility (1 = *very implausible*, 7 = *very plausible*), similarity between the actual memory and counterfactual event (1 = *very dissimilar*, 7 = *very similar*), and novelty (1 = *absolutely sure not novel*, 1 = *absolutely sure novel*). All participants were instructed that judgments of similarity must be based upon how similar the actual memory is to the imagined situation in which the counterfactual event took place. The experimenter also explained to each participant that the novelty rating was based upon whether the participant had previously simulated the counterfactual event outside of the laboratory prior to session 2. Accordingly, participants were told that completely novel counterfactuals have not been simulated outside of the laboratory, while counterfactual that are *not* novel have been simulated outside of the laboratory. The order of these phenomenological ratings was random. Responses faster than 200 ms were excluded from statistical analyses. Postexperimental interviews indicated that no participants were aware of the real purpose of the study. At the end, participants were debriefed and monetarily compensated for their time.

4. Results

Means and standard deviations for phenomenological ratings are presented in Table 1, split by memory valence (positive or negative) and counterfactual direction (upward or downward).

4.1. Effects of direction and valence

We first conducted four separate (Bonferroni-adjusted $\alpha = 0.0125$) two-way ANOVAS to investigate whether counterfactual direction (upward or downward) and/or memory valence (positive or negative) have an effect on counterfactual plausibility, perceived similarity, novelty, and/or ease of counterfactual simulation. The ANOVAS testing the effect of memory valence and counterfactual direction on counterfactual plausibility and ease of simulation, respectively, revealed no significant main effects or interactions (all p's > .32). An ANOVA testing the effect of memory valence and counterfactual direction on perceived similarity revealed a small but significant main effect of memory valence, $F(1, 1{,}289) = 10.123$, $p = .001$, $\eta_p^2 = 0.008$ $\eta_p^2 = 0.008$, and a small but significant main effect of counterfactual direction, $F(1, 1{,}289) = 21.462$, $p < .001$, $\eta_p^2 = 0.016$. These main effects will not be discussed further, however, because there was also a significant interaction between memory valence and counterfactual direction, $F(1, 1{,}289) = 74.639$,

Table 1
Mean (*SD*) phenomenological ratings for upward and downward episodic counterfactual simulations are presented, split by the valence of the memory from which the counterfactual was generated

	Positive Memories		Negative Memories	
Ratings	Upward	Downward	Upward	Downward
Plausibility	4.31 (1.98)	4.23 (1.92)	4.17 (1.91)	4.31 (1.87)
Similarity	3.99 (1.77)	2.72 (1.70)	2.86 (1.70)	3.25 (1.69)
Memory valence	5.78 (0.99)	5.65 (0.96)	2.18 (0.82)	2.32 (0.88)
Counterfactual valence	5.98 (1.03)	2.14 (1.12)	5.28 (1.44)	1.77 (1.13)
Novelty	3.82 (1.40)	4.18 (1.25)	3.70 (1.43)	3.87 (1.40)
Ease	4.86 (1.76)	4.74 (1.65)	4.74 (1.79)	4.68 (1.85)

Note. Mean (*SD*) ratings for each measure were computed from a 7-point Likert scale.

$p < .001$, $\eta_p^2 = 0.055$. Subsequent tests of simple main effects were performed. The simple main effect of counterfactual direction within negative memories was significant, $F(1, 658) = 8.387$, $p = .004$, $\eta_p^2 = 0.013$, such that downward counterfactuals generated from negative memories were perceived as more similar to their corresponding memories than upward counterfactuals generated from negative memories. There was also a simple main effect of counterfactual direction within positive memories, $F(1, 631) = 84.289$, $p < .001$, $\eta_p^2 = 0.118$, such that upward counterfactuals generated from positive memories were perceived as more similar to their corresponding memories than downward counterfactuals generated from positive memories. The final ANOVA testing the effect of memory valence and counterfactual direction on counterfactual novelty revealed only a small but significant main effect of counterfactual direction, $F(1, 1{,}202) = 11.865$, $p = .001$, $\eta_p^2 = 0.010$, such that downward counterfactuals ($M = 4.025$, $SD = 1.332$) were more novel than upward counterfactuals ($M = 3.755$, $SD = 1.416$).

4.2. Correlational analyses between plausibility and similarity, novelty and ease

Bivariate correlational analyses split by counterfactual direction and memory valence indicated that perceived similarity was significantly and positively correlated with plausibility for downward counterfactuals generated from negative memories, $r(333) = 0.462$, $p < .001$, upward counterfactuals generated from negative memories, $r(322) = 0.331$, $p < .001$, downward counterfactuals generated from positive memories, $r(322) = 0.376$, $p < .001$, and upward counterfactuals generated from positive memories, $r(314) = 0.511$, $p < .001$. Additional bivariate correlational analyses split by counterfactual direction and memory valence indicated that counterfactual novelty was only significantly correlated with plausibility for upward counterfactuals generated from negative memories, $r(300) = -0.261$, $p < .001$, and upward counterfactuals generated from positive memories, $r(292) = -0.274$, $p < .001$. There was no significant relationship between counterfactual novelty and plausibility for downward counterfactuals generated from negative memories, $r(309) = -0.067$, $p > .23$, or downward counterfactuals generated from positive memories, $r(297) = -0.038$, $p > .51$. Furthermore, bivariate correlational analyses split by counterfactual direction and memory valence indicated that ease of counterfactual imagining was significantly and positively correlated with plausibility for downward counterfactuals generated from negative memories, $r(329) = 0.185$, $p = .001$, upward counterfactuals generated from negative memories, $r(323) = 0.191$, $p = .001$, downward counterfactuals generated from positive memories, $r(318) = 0.110$, $p = .050$, and upward counterfactuals generated from positive memories, $r(313) = 0.241$, $p < .001$. All correlational results are summarized in Table 2.

Because these bivariate correlational analyses split by counterfactual direction and memory valence indicated that there were only asymmetries in correlational patterns for counterfactual direction (counterfactual novelty and plausibility were significantly related for upward but not downward counterfactual events), we collapsed across positive and negative memories for all subsequent analyses. Steiger's Z-test for dependent correlations was performed to identify potential differences in the magnitude of identified significant

Table 2
Summary of all correlations (uncorrected for multiple comparisons) between all four variables of interest, split by memory valence (positive or negative) from which the counterfactual was generated and counterfactual direction (upward or downward)

	Positive Memories				Negative Memories			
Variables	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Upward								
(1) Plausibility	–				–			
(2) Similarity	0.511***	–			0.331***	–		
(3) Novelty	−0.274***	−0.090	–		−0.261***	−0.108†	–	
(4) Ease	0.241***	0.238***	−0.215***	–	0.191**	0.129*	−0.073	–
Downward								
(1) Plausibility	–				–			
(2) Similarity	0.373***	–			0.465***	–		
(3) Novelty	−0.038	−0.181**	–		−0.067	−0.023	–	
(4) Ease	0.114*	0.090	0.001	–	0.181**	0.142**	−0.085	–

Note. ***$p < .001$, **$p < .01$, *$p < .05$, †$p < .10$.

correlations (Steiger, 1980). For downward counterfactuals, the magnitude of the correlation coefficient between similarity and plausibility ($r(651) = 0.417$, $p < .001$) was significantly greater than the magnitude of the correlation coefficient between ease and plausibility ($r(649) = 0.149$, $p < .001$; Steiger's $Z = 5.525$, $p < .001$). For upward counterfactuals, the magnitude of the correlation between similarity and plausibility ($r(638) = 0.413$, $p < .001$) was significantly greater than both the magnitude of the correlation coefficient between novelty and plausibility ($r(594) = -0.266$, $p < .001$; Steiger's $Z = 2.990$, $p = .003$) and the magnitude of the correlation coefficient between ease and plausibility ($r(638) = 0.217$, $p < .001$; Steiger's $Z = 4.208$, $p < .001$).

4.3. Model fitting results

Using model fit procedures within a linear multiple regression framework, we next sought to determine which combination (if any) of variables—similarity, novelty, and ease of imagining—best accounted for variance in perceived similarity. For downward counterfactual events, after completing backward/forward stepwise regression using Mallow's C_p (Gilmour, 1996; Mallows, 1973), the model that emerged was significant, $F(2, 646) = 72.01$, $p < .001$, Adjusted $R^2 = 0.18$, and included two predictors: similarity ($b = 0.446$, $SE = 0.040$, $\beta = 0.403$, $p < .001$, 95% CI [0.368, 0.523], VIF = 1.01), and ease of imagining ($b = 0.112$, $SE = 0.039$, $\beta = 0.103$, $p = .004$, 95% CI [0.036, 0.188], VIF = 1.01). For upward counterfactual events, after completing backward/forward stepwise regression using Mallow's C_p, the model that emerged was significant, $F(3, 588) = 61.66$, $p < .001$, Adjusted $R^2 = 0.24$, and included all three predictors: similarity ($b = 0.407$, $SE = 0.039$, $\beta = 0.380$, $p < .001$, 95% CI [0.330, 0.484], VIF = 1.04), novelty ($b = -0.304$, $SE = 0.051$, $\beta = -0.217$, $p < .001$, 95% CI [−0.404, −0.204], VIF = 1.02), and ease of imagining (b = 0.119, SE = 0.041, β = 0.108, p = .004, 95% CI [0.039, 0.200], VIF = 1.06). Evaluation of condition indices and variance inflation factors showed no issues with multicollinearity for either model. Evaluation of residual plots showed that the conditions of linearity, normality, and constant variance were satisfied for both models. Collectively, these results show that regardless of counterfactual direction, similarity and ease of imagining are both strong and independent predictors of counterfactual plausibility; but for upward counterfactuals, all three variables—similarity, novelty, and ease of imagining—were significant predictors in the final model and captured unique variance in perceived plausibility. Nevertheless, in both models, effect size estimates show that similarity best explained the variance in counterfactual plausibility.

5. Discussion

In this study, we investigated comparative similarity between actual and counterfactual events as a factor influencing people's judgments of plausibility in counterfactual simulations. We also compared similarity between actual and counterfactual events against other factors thought to influence judgments of plausibility. These analyses yielded five main findings. First, our results show that the greater the perceived similarity between the original memory and the episodic counterfactual event, the greater the perceived plausibility of that counterfactual event, regardless of the emotional content of the memory or the direction (upward or downward) of the counterfactual. While this relationship was significant regardless of memory valence or counterfactual direction, the magnitude of the effect was largest when participants simulated a better alternative to an already positive memory or a worse alternative to an already negative memory. Second, upward episodic counterfactual events that had previously been simulated outside of the laboratory setting (i.e., counterfactual events that were not novel) were perceived as less plausible than truly novel episodic counterfactual events. However, there was no significant relationship between counterfactual plausibility and counterfactual novelty for downward episodic counterfactual events. Third, consistent with prior research on the simulation heuristic (Kahneman & Miller, 1986; Kahneman & Tversky, 1982), we show that enhanced ease of simulation of episodic counterfactual thoughts was associated with greater perceived counterfactual plausibility. Fourth, for both upward and downward counterfactuals, the magnitude of the correlation coefficient between similarity and plausibility was significantly greater than the magnitude of the correlation coefficients between not only ease of imagining and plausibility but also counterfactual novelty and plausibility. Fifth, model fit procedures conducted in a multiple regression framework showed that regardless of counterfactual direction, similarity and ease of imagining were both significantly related to counterfactual plausibility and both captured unique variance in counterfactual plausibility; but for upward counterfactuals, all three variables of interest—similarity, novelty, and ease of imagining—were significantly and uniquely related to plausibility in the final model. Nevertheless, effect size estimates showed that similarity was most closely related to plausibility.

Philosopher David Lewis, in his seminal work *Counterfactuals* (1973; see also Lewis, 1979), suggested that when we evaluate the truth of a counterfactual statement—for example, had A occurred, then B would have occurred—we imagine a possible world in which A is the case (i.e., an A-world). If in this A-world B is also the case and this A-world is comparatively similar to the actual world in which A does not obtain, then we judge the counterfactual as being relatively more likely to have occurred than in an alternative A-world that is relatively less similar to our actual world in which B also obtains. Lewis introduced his notion of comparative similarity to offer a method for determining the semantic value of counterfactual statements. However, we maintain that a version of Lewis's notion of comparative similarity can be usefully investigated as a psychological construct using empirical methods. We put forth the idea that individuals, as a matter of psychological fact, may employ something like Lewis's suggested comparative similarity when evaluating the perceived plausibility of counterfactual events. Thus, inspired by the conceptual framework afforded by Lewis (1973, 1979), we sought to empirically investigate whether variance in judgments of plausibility for imagined counterfactual events could be explained by appealing to comparative similarity judgments of the closeness between the actual world and an imagined possible world in which the counterfactual event obtains. Our results showed that not only was the judged similarity between the actual memory and the counterfactual event significantly related to counterfactual plausibility, but also counterfactual plausibility exhibited a stronger relationship with judgments of similarity than it did with ease of imagining and counterfactual novelty. This was the case regardless of counterfactual direction or the valence of the memory from which the counterfactual was generated.

Furthermore, although the effect sizes were only small or moderate, same direction counterfactual simulations (i.e., better alternative to an already positive memory or a worse alternative to an already negative memory) were judged as more similar to their corresponding memories than opposite direction counterfactual simulations (i.e., worse alternative to a positive memory or a better alternative to a negative memory). The magnitude of the significant correlations between similarity and plausibility for same direction counterfactuals were also larger than those of opposite direction counterfactuals. While this might suggest that same direction counterfactual simulations tend to be anchored on their corresponding memories and, therefore, diverge less from those memories, similarity and plausibility were still strongly correlated for both same and opposite direction counterfactual simulations.

The simulation heuristic provides one means for explaining patterns of judgments about counterfactual events. The simulation heuristic refers to the tendency for people to determine the subjective plausibility of a counterfactual event having occurred based on the ease with which the counterfactual event is imagined. Kahneman and Tversky (1982) argued that the easier it is to imagine the counterfactual event having occurred, the more plausible the counterfactual event seems. Most studies on counterfactual thinking have focused on the simulation of counterfactual alternatives to impersonal events and/or decision making tasks confined to lab settings, and few have employed stimuli extracted from the participant's own episodic autobiographical recollections. Because the majority of these past studies have relied heavily on measures of how participants undo the outcome of some fictional story, little is known about cognitive consequences of episodic counterfactual thinking (De Brigard & Giovanello, 2012; De Brigard et al., 2013). Using episodic autobiographical recollections and episodic counterfactual events, we show for the first time that increased ease of simulation is significantly related to higher ratings of perceived counterfactual plausibility. This finding corroborates and expands upon existing work on the simulation heuristic (Kahneman & Miller, 1986; Kahneman & Tversky, 1982). However, we also show that ease of imagining is not the only factor influencing ratings of plausibility, nor is it the factor that best accounts for the increase in perceived plausibility.

Recent evidence has suggested that, regardless of counterfactual direction, repeatedly simulating the same counterfactual events in the laboratory reduces their novelty, rendering those counterfactual events less plausible (De Brigard et al., 2013). Instead of manipulating repetition in a laboratory setting to create counterfactuals that are not novel, we directly measured counterfactual novelty by asking participants if they had previously thought about each counterfactual event outside of the laboratory setting and how confident they were in each of these judgments. De Brigard et al. (2013) interpreted their results within the mental-models framework on counterfactual generation (Byrne, 1997, 2002; Johnson-Laird & Byrne, 2002). Following the mental-models view on episodic counterfactual simulation, if the initially generated counterfactual diverges minimally from the actual memory, the perceived plausibility of the counterfactual event should be high. But when more attention can be allocated to increase the detail of the counterfactual event (e.g., through repeated simulation), the counterfactual event further diverges from the actual memory, which renders the counterfactual event less plausible.

Partially corroborating findings from De Brigard et al. (2013), this study revealed that upward counterfactual events that participants simulated outside the laboratory were perceived as less plausible than counterfactual events participants had not thought about before the experimental session. However, unlike De Brigard et al. (2013), we did not identify a significant relationship between counterfactual plausibility and counterfactual novelty for downward counterfactual events. We believe this discrepancy can be accounted for, however. Considerable evidence has accumulated showing that the simulation of upward counterfactuals is accompanied by negative emotions, such as regret, shame, sadness, or distress (Davis, Lehman, Wortman, Silver, & Thompson, 1995; Epstude & Jonas, 2015; Gilbar & Hevroni, 2007; Mandel, 2003; Niedenthal, Tangney, & Gavanski, 1994; Zeelenberg et al., 1998), whereas the simulation of downward counterfactuals is accompanied by positive emotions, such as contentment, satisfaction, or relief (Byrne, 2016; Guttentag & Ferrell, 2004; Johnson, 1986; Johnson & Sherman, 1990; Markman, Gavanski, Sherman, & McMullen, 1993). Because an important and pervasive function of counterfactual thinking is to alleviate negative emotion and serve a mood repair function, often through inducing self-serving biases (Byrne, 2016; Epstude & Roese, 2008; Koehler, 1991; Roese, 1994), repeated simulation of counterfactual events may modulate their perceived plausibility to better fulfill these functions. If one repeatedly simulates a better alternative event and comes to believe that the event is less and less plausible with repeated simulation, this may actually alleviate negative affect or even induce positive affect. For if one believes that a better event could not have plausibly occurred, then negative feelings of regret, shame, sadness, or distress that are typically associated with upward counterfactual thinking are less likely to accompany the simulation. Therefore, we may naturally and repeatedly simulate upward counterfactual events (i.e., render them less novel) and perceive them as less plausible with repeated simulation to avoid or alleviate negative mood and/or induce positive feelings. Whether this is, in fact, the correct explanation for these findings remains an open question, and one that merits further research.

In sum, the present work suggests that judgments of comparative similarity between actual memories and counterfactual events, ease of imagining counterfactual events, and counterfactual novelty all exhibit strong and independent relationships with counterfactual plausibility, but comparative similarity best explains the variance in judgments of counterfactual plausibility. It is worth noting, however, that there is still a considerable amount of remaining variance to be explained in judgments of counterfactual plausibility. Future work will seek to determine whether a greater proportion of variance in counterfactual plausibility can be explained by investigating only controllable and unstable features of a situation while excluding uncontrollable and stable features (Girotto, Legrenzi, & Rizzo, 1991; Kahneman & Miller, 1986). The value in understanding why certain kinds of counterfactual events are considered more or less plausible is underscored by the wide-ranging effects that assessments of counterfactual plausibility have upon judgments of regret, causation, affect, and responsibility (Petrocelli, Percy, Sherman, & Tormala, 2011).

References

Bowie, G. L. (1979). The similarity approach to counterfactuals: Some problems. *Nous*, *13*(4), 477–498.

Byrne, R. M. J. (1997). Cognitive processes in counterfactual thinking about what might have been. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 37, pp. 105–154). San Diego, CA: Academic Press.

Byrne, R. M. J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences*, *6*, 426–431. doi:10.1016/S1364-6613(02)01974-5

Byrne, R. M. J. (2005). *The rational imagination*. Cambridge, MA: MIT Press.

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157. doi:10.1146/annurev-psych-122414-033249

Davis, C. G., Lehman, D. R., Wortman, C. B., Silver, R. C., & Thompson, S. C. (1995). The undoing of traumatic life events. *Personality and Social Psychology Bulletin*, *21*, 109–124. doi:10.1177/0146167295212002

De Brigard, F., & Giovanello, K. S. (2012). Influence of outcome valence in the subjective experience of episodic past, future and counterfactual thinking. *Consciousness and Cognition*, *21*, 1085–1096. doi:10.1016/j.concog.2012.06.007

De Brigard, F., Szpunar, K. K., & Schacter, D. L. (2013). Coming to grips with reality: Effect of repeated simulation on the perceived plausibility of episodic counterfactual thoughts. *Psychological Science*, *24*, 1329–1334. doi:10.1177/0956797612468163

Epstude, K., & Jonas, K. J. (2015). Regret and counterfactual thinking in the face of inevitability: The case of HIV positive men. *Social Psychological and Personality Science*, *6*, 157–163. doi:10.1177/1948550614546048

Epstude, K., & Roese, N. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, *12*, 168–192. doi:10.1177/1088868308316091

Fine, K. (1975). Review of David Lewis's *Counterfactuals*. *Mind*, *84*, 451–458.

Gilbar, O., & Hevroni, A. (2007). Counterfactuals, coping strategies and psychological distress among breast cancer patients. *Anxiety, Stress, & Coping*, *20*, 383–392. doi:10.1080/10615800701384439

Gilmour, S. G. (1996). The interpretation of Mallow's C_p –statistic. *Journal of the Royal Statistical Society, Series D*, *45*, 49–56. doi:10.2307/2348411

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in causal thinking. *Acta Psychologia*, *78*, 111–133. doi:10.1016/0001-6918(91)90007-M

Guttentag, R., & Ferrell, J. (2004). Reality compared with its alternatives: Age differences in judgments of regret and relief. *Developmental Psychology*, *40*, 764–775. doi:10.1037/0012-1649.40.5.764

Johnson, J. T. (1986). The knowledge of what might have been: Affective and attributional consequences of near outcomes. *Personality and Social Psychology Bulletin*, *12*, 51–62. doi:10.1177/0146167286121006

Johnson, M. K., & Sherman, S. J. (1990). Constructing and reconstructing the past and future in the present. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 482–526). New York: Guilford Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678. doi:10.1037/0033-295X.109.4.646

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153. doi:10.1037/0033-295X.93.2.136

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, *110*, 499–519. doi:10.1037/0033-2909.110.3.499

Lewis, D. K. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Lewis, D. K. (1979). Counterfactual dependence and time's arrow. *Nous*, *13*, 455–476. doi:10.2307/2215339

Mallows, C. L. (1973). Some comments on C_p. *Technometrics*, *15*, 661–675. doi:10.2307/1267380

Mandel, D. (2003). Counterfactuals, emotions, and context. *Cognition and Emotion*, *17*, 139–159. doi:10.1080/02699930302275

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, *29*, 87–109. doi:10.1006/jesp.1993.1005

Niedenthal, P. M., Tangney, J. P., & Gavanski, L. (1994). "If only I weren't" versus "If only I hadn't": Distinguishing shame and guilt in counterfactual thinking. *Journal of Personality and Social Psychology*, *67*, 585–595. doi:10.1037/0022-3514.67.4.585

Petrocelli, J. V., Percy, E. J., Sherman, S. J., & Tormala, Z. L. (2011). Counterfactual potency. *Journal of Personality and Social Psychology*, *100*, 30–46. doi:10.1037/a0021523

Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, *66*, 805–818. doi:10.1037/0022-3514.66.5.805

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245–251. doi:10.1037/0033-2909.87.2.245

Szpunar, K. K., & Schacter, D. L. (2013). Get real: Effects of repeated simulation and emotion on the perceived plausibility of future experiences. *Journal of Experimental Psychology: General*, *142*, 323–327.

Williamson, T. (1988). First-order logics for comparative similarity. *Notre Dame Journal of Formal Logic*, *29*(4), 457–481.

Zeelenberg, M., van Dijk, W., van der Pligt, J., Manstead, A., van Empelen, P., & Reinderman, D. (1998). Emotional reactions to the outcomes of decisions: The role of counterfactual thought in the experience of regret and disappointment. *Organizational Behavior and Human Decision Processes*, *75*, 117–141. doi:10.1006/obhd.1998.2784