Do episodic counterfactual thoughts focus on personally controllable action?

Roese, Neal J.; Smallman, Rachel; Epstude, Kai

Published in:
Journal of Experimental Social Psychology

DOI:
10.1016/j.jesp.2017.05.006

Publication date:
2017

Journal of Experimental Social Psychology 73 (2017) 14–23

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Do episodic counterfactual thoughts focus on controllable action?: The role of self-initiation

Neal J. Roese[a,*], Rachel Smallman[b], Kai Epstude[c]

[a] Northwestern University, United States
[b] Texas A & M University, United States
[c] University of Groningen, Netherlands

ABSTRACT

Counterfactual thoughts refer to alternatives to the past. Episodic counterfactual thoughts have in past research been shown to be primarily goal-directed and to engender performance improvement. Some past research supports this perspective with the observation that episodic counterfactuals center mostly on controllable action, whereas other research does not show this. We offer a theoretical resolution for these discrepant findings centering on the role of self-initiation, such that counterfactuals more often focus on internally controllable action to the extent that the circumstance is one that was self-initiated rather than initiated by others. In doing so, we disambiguate two dimensions of causal explanation: locus (self vs. other) and controllability (high vs. low) that previous studies conflated, demonstrating that variation as a function of self-initiation in the content of episodic counterfactuals occurs primarily along the former but not the latter dimension. These results support the functional theory of counterfactual thinking.

1. Introduction

Counterfactual thinking refers to thoughts about what might have been, of how the past might have been different had some aspect been different (Byrne, 2016; Roese, 1997). Counterfactuals may be understood as instantiations of conditional propositions, containing an antecedent ("if") and consequent ("then"), as in "If only I had studied, then I would have passed the exam." Echoing recent contributions (e.g., De Brigard & Giovanello, 2012; Özbek, Bohn, & Berntsen, 2016; Schacter, Benoit, De Brigard, & Szpunar, 2015), we distinguish episodic counterfactuals from semantic counterfactuals, such that the former focus on personally meaningful alternatives to events that were experienced first-hand (as in the example of writing an exam), whereas the latter focus on alternative constructions derived from general knowledge of history, society, and the natural world (e.g., Quelhas & Byrne, 2003; Revlin, Cate, & Rouss, 2001; Thompson & Byrne, 2002). A prominent view of episodic counterfactual thinking is that they primarily serve a preparative function, which is to say that they contain insights as to how an alternative past action might have brought about goal success, which then feeds into subsequent action that facilitates goal success (Epstude & Roese, 2008; Roese & Epstude, 2017).

Recently, a discrepancy has emerged in terms of whether counterfactuals do or do not center mainly on internally controllable action. That episodic counterfactuals center mostly on controllable action is a key tenet of the functional theory of counterfactual thinking. As we review below, a substantial number of earlier studies supported this idea, but newer evidence contradicts it. The present research aims to account for this variability across research reports. We propose and demonstrate that episodic counterfactuals are more likely to focus on internally controllable action to the extent that the situation in question is self-initiated as opposed to other-initiated. We consider the theoretical basis for this contention in the next paragraphs.

1.1. The functional theory of counterfactual thinking

The operation of episodic counterfactual thinking may be usefully illuminated by the functional theory of counterfactual thinking (Epstude & Roese, 2008, 2011; Roese, 1994, 1997, 1999; Roese & Epstude, 2017), which seeks to describe observable patterns in terms of goal cognition. Episodic counterfactuals usually embody goals and specify means by which those goals may have been achieved. Counterfactuals relate directly to planning and action implementation, which may in turn guide behavior (i.e., a preparative function). Counterfactual thoughts may bring about performance improvement via either a content-specific pathway (in which the counterfactual

* Corresponding author at: Kellogg School of Management, Northwestern University, 2211 Campus Drive, Evanston, IL 60208, United States.
E-mail address: n-roese@kellogg.northwestern.edu (N.J. Roese).

http://dx.doi.org/10.1016/j.jesp.2017.05.006
Received 27 January 2017; Received in revised form 24 May 2017; Accepted 24 May 2017
Available online 06 June 2017

insight directly informs behavior change by specifying the particular means) or a content-neutral pathway (in which counterfactual thought activates cognitive procedures or negative affect that bring about behavior change independently of the specific informational content of the counterfactual itself).

One principle of the theory is that form fits function, which is to say that the form (e.g., direction or content) of a counterfactual thought will vary in terms of how useful it is for goal progress, such that those counterfactual forms most amenable to behavior change will become more numerous under circumstances in which performance improvement is possible. In particular, counterfactuals that focus on personally controllable action constitute a form that is better suited to serving the function of goal progress because only personally controllable actions can be deployed by the individual in the service of goal striving. In observing that counterfactual thoughts for the most part center on personally controllable action, the theoretical insight is that episodic counterfactuals primarily serve a preparative function (although other functions may also be served, albeit less frequently).

It is important to clarify that in recent writings, the term controllable counterfactual has been used as a generic category that conflates two distinct dimensions of 1) locus of causation (internal vs. external) and 2) controllability (controllable vs. uncontrollable) that were previously specified by causal attribution theorists, particularly Weiner (1985, 1986). In Weiner's view, the content of lay causal attributions can be characterized by three orthogonal dimensions (the third being stability), and that particular attributions differentially positioned along these dimensions will have different effects on emotion and motivation. Because counterfactuals imply causal relations (as captured by their if-then conditional structure), applying the past insights of attribution theorists may help to illuminate patterns of counterfactual thinking. Accordingly, one contribution of the present research is to specify whether episodic counterfactuals vary meaningfully in terms of all three dimensions of locus, controllability, and stability. Moving forward, we will use the term “internally controllable” to refer generically to current or past studies that do not distinguish the locus and controllability dimensions, or else we will specify the relevant dimension using Weiner's terminology.

1.2. Evidence that counterfactuals center mainly on internally controllable factors

Supporting the functional theory of counterfactual thinking, much early research indicated that counterfactuals centered mainly on internal and controllable factors. For example, Mandel and Lehman (1996) presented participants with a scenario describing an automobile accident with various aspects of the episode varying in the degree to which they could be controlled by the focal actor. Participants' counterfactual responses to the scenario tended to alter the more controllable of those aspects. Girotto, Legrenzi, and Rizzo (1991) presented participants with a scenario describing a medical mishap along with a variety of antecedents that varied in their controllability by the protagonist. Counterfactual thoughts (collected in response to an “if only” prompt) more often focused on controllable than uncontrollable antecedents. McCloy and Byrne (2000) used a modification of the Girotto et al. (1991) scenario and reported similar results. McEleney and Byrne (2006) used a scenario along with a diary-creation dependent measure from which instances of counterfactual thinking were coded; participants produced more counterfactual thoughts for controllable than uncontrollable outcomes. Morris, Moore, and Sim (1999) presented participants with a scenario involving an industrial accident and found that resulting counterfactuals typically undid the accident by focusing on human error rather than systemic or organizational factors. Roese and Olson (1995) and Rye, Cahoon, Ali, and Daftary (2008) manipulated controllability in a scenario and found that more counterfactuals followed from controllable than uncontrollable outcomes.

Thus, several scenario studies showed that counterfactual thoughts tend to focus on personally controllable actions. However, subject speculations from scenario studies are not the same as episodic counterfactuals, which are better examined via self-reports of experienced episodes, either recalled from daily life or reported in light of laboratory tasks. It is to this type of evidence that we turn next.

Mandel (2003) asked participants to recall negative experiences, manipulated to focus either on an academic or interpersonal event. Participants then reported whether they had had any counterfactual thoughts, and if so, details about them. These reported counterfactuals tended overall to be internal (i.e., self-focused vs. other-focused). Participants also provided a scale rating of perceived control; availability of self-focused counterfactuals correlated significantly with controllability ratings. Further, Davis, Lehman, Wortman, Silver, & Thompson (1995, Study 2) conducted interviews among bereaved parents who had lost their infant to Sudden Infant Death Syndrome. Counterfactual thinking was common among these parents and, pivotally, 90% of their counterfactuals specified actions they personally could have taken (or not taken) to have prevented their tragic loss. Finally, Epstude and Jonas (2015) found that HIV + individuals who reported counterfactuals about their infection also indicated having more control over their infection compared to those who did not report having counterfactuals.

As an example of a study using a laboratory task to assess counterfactuals as they occur in response to an evoking episode, Hammell and Chan (2016) had their participants play Nintendo Wii console video games. Counterfactual thinking was prompted and then coded by raters; counterfactuals focusing on controllable aspects outnumbered those focusing on uncontrollable aspects by a factor of 2 to 1 (however, the underpowered statistical tests were non-significant). Two decades earlier, Markman, Gavanski, Sherman, and McMullen (1995) used a gambling task in which participants played a computer-simulated “wheel of fortune” game, with their degree of control over the task experimentally manipulated to involve either control of which of two wheels dictated their payoff or control of the stopping point of the wheel. The game was fixed so that payoffs were constant yet would generate varying degrees of counterfactual thinking based on how close the wheel came to a larger payoff. Counterfactual thinking was assessed via direct prompts; these counterfactuals tended to focus on that aspect of the game (wheel choice vs. stopping point choice) over which participants believed they had controlled.

Thus, substantial evidence accumulated over twenty years of published research suggests that counterfactual thinking is most likely to focus on factors that are internal (vs. external) and controllable (vs. uncontrollable). Next, we turn to evidence to the contrary, and here the evidence is almost entirely based on laboratory experiences.

1.3. Evidence that counterfactuals do not center mainly on internally controllable factors

Girotto, Ferrante, Pighin, and Gonzalez (2007) described eight experiments that each manipulated whether participants experienced first-hand versus read about a laboratory outcome centering on task performance. The procedure typically involved a blind choice of task selection, task completion, and then bogus failure feedback. The dependent variable was counterfactual thinking, prompted to focus on the upward direction of comparison. These authors drew the conclusion that counterfactuals tended to “alter uncontrollable events … rather than controllable ones” (p. 515); however, no direct contrasts between proportions of counterfactuals centering on more versus less controllable aspects were presented. Pighin, Byrne, Ferrante, Gonzalez, and Girotto (2011) did provide these contrasts and showed that participants who directly experienced the outcome reported fewer counterfactuals centering on their controllable actions than did those who read about the experience. This research therefore captures episodic counterfactual thoughts, and suggests a possible difference between results from scenario studies (which produce greater focus on controllable actions)

versus laboratory studies (which produce lesser focus on controllable action), mirroring widely reported differences in people's affective forecasts versus actual experiences (e.g., Wilson & Gilbert, 2003). We will return to this possibility in the next section.

Ferrante, Girotto, Stragà, and Walsh (2013, Study 1) used a method involving an anagram task with accurate feedback, after which participants provided upward counterfactuals in response to a prompt. The counterfactuals, coded by an independent rater, showed that a relative minority (25%) centered on controllable aspects (e.g., concentration, attention, reasoning tactics, etc.) versus uncontrollable (e.g., state, traits, abilities, situational aspects, etc.). In Study 2 of that paper, in which participants had a choice over whether the task was easy or difficult (only choosers of the difficult task, a minority, were analyzed), resulted in a somewhat higher preponderance of counterfactuals focusing on controllable means (43%) presumably because task choice offered an additional controllable target on which participants might focus. Ferrante, Stragà, Walsh, and Girotto (2016) reported a similar laboratory task in which the yield of controllable (vs. uncontrollable) counterfactuals was 25%. Intriguingly, in that same paper, a study of marathon runners reflecting on their race performance revealed a very different rate of 68% of counterfactuals focusing on internally controllable actions. By contrast, Mercier et al. (2016), using laboratory tasks involving word search or syllogisms, reported rates of counterfactuals that averaged 16% across studies. An aspect of all three of these papers has been to benchmark the rate of internally controllable counterfactuals against the rate of controllable prefactuals (i.e., future-focused if-then contingencies, Epstude, Scholl, & Roese, 2016); the general finding was that the latter are more likely to focus on controllable actions than the former.

Thus, in contrast to the evidence reported in the previous section, these more recent studies report that counterfactual thinking tends not to focus on internal or controllable factors (See Table 1 for summary). These new results, as pointed out by Mercier et al. (2016), clearly call into question a key tenet of the functional theory of counterfactual thinking. How possibly can counterfactuals provide insights for personal improvement if they do not focus on internally controllable actions? The theoretical puzzle is twofold: first, how best to reconcile the discrepancy between those earlier studies showing that counterfactuals focus on internal and controllable factors versus those later studies that do not, and second, how best to explain the considerable variability in the proportion of controllable counterfactuals observed across recent studies (ranging from 9% in Mercier et al., 2016, Study 1b to 68% in Ferrante et al., 2016, Study 2).

1.4. A theoretical resolution

In trying to capture the reason for the discrepancy between the two bodies of findings, a good starting point is to consider the goal structure of the most discrepant findings. At one extreme, we have Mercier et al. (2016, Study 1b), in which only 8% of the counterfactuals focused on internally controllable action. In this experiment, the laboratory task was to order 5 people by height on the basis of 4 comparative statements. At the other extreme, there are the marathon runners who, in reflecting on their race performance, generated counterfactuals that focused on their own actions 68% of the time (Ferrante et al., 2016, Study 2). How do these situations differ? We considered and tested three ways that the situations differ: 1) the normality of the task (an unusual laboratory task with arcane rules is more novel and less normal than a marathon for which one has trained repeatedly), 2) the degree of expertise (e.g., knowledge, training) required to complete the task (one has no expertise in an arcane laboratory task but training for a marathon confers expertise), and 3) self-initiation (laboratory participants are pulled into a task by the experimenter whereas marathon runners voluntarily self-selected into the race). We next elaborate on each of the three factors and why each might potentially explain variation across previous experiments.

1.4.1. Normality

Situations may differ as to whether they are normal versus abnormal (Bear & Knobe, 2016). Normality is the central construct of norm theory (Kahneman & Miller, 1986), which posits that an abnormal event (defined as an antecedent action or situation) is one that deviates from past patterns and elicits counterfactual thoughts that essentially recapitulate the past pattern. That is, if an unusual action is taken, the counterfactual focuses on an alternative in which that unusual action is replaced by the normal action (Gavanski & Wells, 1989; Kahneman & Tversky, 1982; Miller & McFarland, 1986; Roese & Hur, 1997). For example, if an individual finds a $20 bill on the sidewalk while walking home from work, she might, to the extent that she had taken an abnormal route home, be more likely to consider the counterfactual, “If I hadn't taken this unusual route, I wouldn't have found the money.” In this way, the counterfactual represents an imagined return to a more normal circumstance (it is normal to take one's usual route home). This pattern fits older attribution research, in which the dimension of stability refers to whether a causal factor is one that tends to change around versus remain constant. Effort and ability are respective examples of causes that are unstable versus stable. On average, causal attributions tend to focus more on unstable (e.g., the decision as

Table 1
Rate of internally controllable counterfactuals in recent and current experiments.

Paper	Counterfactual condition	Benchmark condition	Task	Self-initiated
Ferrante et al. (2013, Study 1)	25%	50% (prefactual)	Scrambled word	No
Ferrante et al. (2013, Study 2)	43%	78% (prefactual)	Syllogism	No
Ferrante et al. (2016, Study 1)	25%	78% (prefactual)	Word search	No
Ferrante et al. (2016, Study 2)	62%	85% (prefactual)	Marathon	Yes
Mercier et al. (2016, Study 1a)	35%	78% (prefactual)	Word search	No
Mercier et al. (2016, Study 1b)	9%	11% (positive outcome)	Syllogism	No
Mercier et al. (2016, Study 2)	15%	72% (advice)	Syllogism	No
Current (Pretest)	71%	53% (positive outcome)	Recall	Yes
Current (Expt 1)	77%	72% (abnormal)	Recall	Yes
Current (Expt 2)	73%	75% (inexpert)	Recall	Yes
Current (Expt 3)	85%	65% (other-initiated)	Recall	Yes
Current (Expt 5)				
Locus	92%	44% (other-initiated)	Recall	Yes
Control	89%	83% (other-initiated)	Recall	Yes
(In)Stability	90%	76% (other-initiated)	Recall	Yes

Note. Experiments are included above that reported proportion data as the dependent measure pertaining to internally controllable counterfactuals. Values are the proportion of counterfactuals reported by participants that focused on internally controllable action. The counterfactual column gives the focal proportion of counterfactual thoughts, whereas the benchmark condition column gives the proportions of various contrast cases, which sometimes were prefactual judgments and other times counterfactuals under varying aspects of the episode, e.g., a positive outcome or an episode initiated by others.

to which route to take) than stable causes (e.g., one's stable trait of openness to experience; see Weiner, 1986, for review of this older research), which is somewhat analogous to the finding that counterfactuals tend to focus on abnormal than normal antecedents. Of key importance, in this prior work, the normality (and stability) dimensions center on (potential) causal antecedents, not the larger episode in which those antecedents (potentially) operate.

In the present research, we wondered whether normality defined at the broader level of the episode itself might make a difference in whether counterfactuals focus on internally controllable action. For example, swimming is normal activity for some but not others. Defined at the level of episode, it seems plausible that normal episodes may bring a host of familiar targets from memory for counterfactuals to focus on. For a habitual swimmer, altering stroke, breathing, or swimming gear (all internally controllable acts) might be salient. For a non-swimmer bumped accidentally into a pool, there would be no such behavioral repertoire in memory to consult, leading perhaps instead to greater emphasis on factors outside of one's control (e.g., "If only the water hadn't been so cold."). Based on these ideas, we tested whether normal episodes are more likely to involve counterfactuals that focus on internally uncontrollable actions, whereas abnormal situations would be less likely to evoke such thoughts.

1.4.2. Expertise

Tasks differ in the extent to which they are amenable to the individual's talents and past experiences. For example, an individual might be highly skilled at tennis but inept at billiards. For experts, with their finer knowledge of task demands and greater insight into their own ability, it is easier to see how their internal and controllable actions might be modified, sometimes subtly, to produce a better performance. By contrast, for novices, the nature and rules of the unfamiliar task demand close attention, with resulting counterfactuals centering more on the task itself rather than on their own controllable actions. Thus, it is possible that internal and controllable counterfactuals are more likely among experts than novices. This facet might explain the discrepancy between research findings in that the more recent laboratory demonstrations involve tasks for which participants are inexpert, thus pushing them away from generating internal and controllable counterfactuals.

1.4.3. Self-initiation

Episodes differ in whether the individual decided deliberately to enter into them versus being lured by others. For example, an individual might go to the cinema because the film is precisely the one that most interests her, or she might instead go because a friend has dragged her along. For a self-initiated episode, the individual can look readily to their own actions, preferences, and knowledge as the basis for generating internal and controllable counterfactual thoughts, in part because of the greater overlap between their own preferences and the nature of the episode. Importantly, we distinguish between factors centering on initiation versus operation of the episode. For example, Janet may initiate a conversation about action movies with her friends Sarah and Julie, but then the conversation may then continue with contributions from all three friends until its termination. Janet's subsequent counterfactual thought may target any of the actions by her or her friends, from the initiation point on. If the factor of self-initiation of an episode pushes individuals toward counterfactuals that are internal and controllable merely because they center on initiation of the episode, the result is largely tautological. But if instead self-initiation additionally pushes people to consider not just initiation but also other aspects of its ongoing operation, then a more general and interesting determinant factor will have been uncovered.

Self-initiation might explain the discrepancy between research findings in that the more recent laboratory demonstrations involve tasks that participants did not think to engage in themselves but rather were asked to engage in by experimenters, thus pushing them away from generating internal and controllable counterfactuals. As noted, the one recent study that produced the greatest proportion of internal and controllable counterfactuals (Ferrante et al., 2016, Study 2) involved runners in a marathon, clearly a circumstance that participants had self-initiated.

We have thus identified three factors – normality, expertise, and self-initiation – that may plausibly explain variation in reported results. In the present research, we manipulated these three factors in successive experiments to specify which best accounted for variability in the rate at which counterfactuals focus on internal and controllable action.

As noted previously, causal attributions may fall along three independent dimensions: locus, controllability, and stability (Weiner, 1985, 1986). This conception provides a useful umbrella for considering the general contours of episodic counterfactual thinking, but in addition, attribution theory research in a previous era provided evidence that is compatible with the functional theory, in that spontaneous causal inferences tend most often to focus on effort and other internal factors (see Weiner, 1986, Tables 2–3, for summary of ten earlier studies along these lines). Weiner cautioned against conflating internal versus external locus of causation with controllability: there can be internal causes that are both controllable (effort) and uncontrollable (physique), and this same distinction applies to others as well as oneself. To clarify our findings and to build a bridge to Weiner's theory, in Experiment 5 we measured not just whether counterfactuals center on controllable action, but the extent to which variation occurred independently across locus, stability, and controllability.

1.5. The present research

We report five experiments that all use the method of retrospective self-reports of personal experiences. One advantage of this method is that it permits direct assessment of episodic counterfactual thoughts (again, those defined in terms of their connection to a personally experienced event). Another advantage is that by permitting participants to recall events from a wide range of life experiences, we gain a broad portrait of cognitions that span discrete types of episodes, thus lending generality to the results.

In all experiments, participants first recalled a personally experienced event, then reported a counterfactual alternative to that event, and finally characterized that counterfactual in one way or another. For Experiments 1–3 the dependent variable was a dichotomous judgment of whether the counterfactual focused on a controllable versus uncontrollable antecedent. Two aspects of this dependent variable bear comment. First, the measure conflates the locus and controllability dimensions, and the decision to do so was made soberly so as to facilitate direct comparison to the most recent crop of relevant published research. Experiment 5 remedies this conflation by unpacking judgments into the three Weinerian dimensions of locus, controllability, and stability. Second, although the controllability judgment may also be captured using a continuous scale rating, we used a dichotomous judgment again to facilitate comparison to recent relevant research. Aiming to generalize beyond the dependent measure that prompts participants for their counterfactual thoughts, Experiment 4 used an open-ended thought-listing task, thus allowing us to capture spontaneously generated counterfactuals. We report all measures, manipulations, and exclusions in the respective sections of each experiment.

2. Pretest

This pretest was a "proof of concept" of the within-subject experimental paradigm. We replicated the widely used manipulation of outcome valence to verify that the paradigm would be sensitive to variation in a factor already known to powerfully impact counterfactual thinking. That is, counterfactual thoughts are more active following negative than positive events (De Brigard & Giovanello, 2012; Hafner, White, & Handley, 2012; Markman, Gavanski, Sherman, & McMullen,

1993; McEleney & Byrne, 2006; Petrocelli & Harris, 2011; Roese & Hur, 1997; Roese & Olson, 1997; Summerville, 2011). Under the contention of the functional theory that episodic counterfactuals aim primarily at goal-striving generally and remedying failure specifically, we would expect that controllable counterfactuals would be more prevalent following failure than success.

Participants reported a salient experience and then recorded one counterfactual alternative to that event. The dependent measure was participants' own dichotomous judgment as to whether the counterfactual centered on a controllable action. Outcome valence was manipulated on a within-subject basis (order randomly counterbalanced).

2.1. Method

The sample comprised 91 adult Americans from Mechanical Turk. The initial sample was $N = 97$; 6 participants were excluded for failure to follow instructions (e.g., failing to record an outcome or a counterfactual either by entering gibberish into the response box or leaving it blank). Of the final sample, 44 were women and 47 were men, with age $M = 34.5$. The survey ran Sept 20–21, 2016.

Participants first recorded an outcome, with outcome valence manipulated on a within-subject basis (order randomized): "Think back to a recent POSITIVE [NEGATIVE] experience. In the space below, please share a few details about this event." Next, participants answered a counterfactual probe, providing open-ended descriptions of counterfactual alternatives to the outcome in question: "After having experiences like this, sometimes people have thoughts like 'what if' ['if only'] in that they can see how things could have gone differently. In the space below, please share one 'what if' ['if only'] thought." For clarity, the phrase "what if" was always used for positive outcomes, whereas the phrase "if only" was always used for negative outcomes. Finally, participants made the dichotomous (yes or no) judgment as to whether the counterfactual embraced a personally controllable action: "Does your 'what if' ['if only'] thought focus on something that you personally could have done? (in other words, something under your own control?)."

2.2. Results

Participants had little difficulty providing examples of positive and negative life outcomes and then providing counterfactuals with regard to those outcomes. Outcome valence significantly affected the rate of controllable counterfactuals, which was greater for negative outcomes (71%) than for positive outcomes (53%), McNemar $\chi^2 = 5.69$, $p = 0.017$, odds ratio = 2.21. To ensure robustness by accounting for potential for order effects, we also analyzed just the first outcome (i.e., positive or negative event) that was presented (randomly) to participants. A similar result emerged from this between-subject analysis, such that there was a higher rate of controllable counterfactuals for negative outcomes (74%) than for positive outcomes (45%), Pearson χ^2 (1, $N = 91$) = 7.78, $p = 0.005$, odds ratio = 0.29.

We wondered whether participants' classification of counterfactuals as internally controllable versus uncontrollable would correspond to those of an external rater. An assistant blind to the research hypotheses coded the counterfactuals as to whether they focused on an internally controllable action. The judgments of this external rater closely corresponded to those of participants, positive outcome $\chi^2(1) = 44.54$, $p < 0.001$; negative outcome $\chi^2(1) = 49.54$, $p < 0.001$. Analysis using external rating judgments instead of participant's own subjective ratings showed the same result, i.e., a higher rate of controllable counterfactuals for negative (69%) than for positive outcomes (46%), McNemar $\chi^2 = 9.76$, $p = 0.002$, odds ratio = 3.10. These results render less plausible the interpretation that the discrepancy involving prior experiments was due to use of participant's own subjective ratings versus those of external raters.

A post hoc power analysis using G*Power 3.1's implementation of the McNemar test indicated 79% obtained power. Moving forward with the same within-subject paradigm, we set sample size for the subsequent experiments on the basis of the pretest power analysis, which indicated that a sample size of $N = 94$ would be needed to achieve 80% power in a McNemar test. Accordingly, we aimed for sample sizes of $N = 100$ for the next experiments.

The key pretest finding is that for a negative outcome, with counterfactuals focusing on a wide range of personal episodes, the rate of controllable counterfactuals was substantial, at nearly three quarters. In replicating a main effect of outcome valence, this pretest establishes the sturdiness of the within-subject retrospective recall paradigm, which was next used to test the three candidate factors of normality, expertise, and self-initiation. Because of the conceptual platform that positions counterfactuals as largely connected to alleviating goal blockage, all subsequent experiments focused exclusively on negative outcomes.

3. Experiment 1

Experiment 1 tested the role of normality. We wondered whether counterfactuals focusing on controllable action might be more prevalent following normal than abnormal episodes, in part because people are accustomed to trouble-shooting difficulties under familiar situations. By contrast, abnormal episodes might push counterfactuals to focus more on less familiar facets of the episode. In this logic, one reason why studies by Ferrante et al. (2013, 2016) and Mercier et al. (2016) obtained low reduced rates of controllable counterfactuals might be that their laboratory procedures are experienced by participants as abnormal, but under more typical life circumstances, controllable counterfactuals may prevail.

3.1. Method

The sample comprised 103 adult Americans from Mechanical Turk, with no exclusions (all followed the instructions correctly and provided legible responses). In this sample, 53 were women and 50 were men, with age $M = 34.8$. The survey ran Oct 6–7, 2016.

The outcome probe was similar to that of the negative outcome probe used in the pretest, but formed the basis of a within-subject manipulation of normality (order randomized): "Think back to a recent negative experience that you also found to be ordinary, typical, and very much [unusual, atypical, and not] a part of your normal routine. In the space below, please share a few details about this event." Next, participants responded to the counterfactual probe and dichotomous counterfactual controllability judgment as per the pretest.

3.2. Results

Normality had no impact on the rate of controllable counterfactuals. Normal episodes evoked controllable counterfactuals at a rate (75%) just as numerous as abnormal episodes (70%), McNemar $\chi^2 = 0.59$, $p = 0.441$, odds ratio = 1.45. As before, we also analyzed just the first outcome (i.e., normal or abnormal events) that participants responded to. Similar results were seen for this between-subject analysis, such that both normal (69%) and abnormal episodes (80%) elicited high rates of controllable counterfactuals, Pearson χ^2 (1, $N = 103$) = 1.72, $p = 0.19$, odds ratio = 0.55.

This experiment thus produced two useful findings: first, normality was not supported as a viable explanation for the discrepancy in rates of controllable counterfactuals, and second, echoing the results of the pretest, we found controllable counterfactuals following a negative, personally meaningful outcome to be quite numerous, at a rate of around three quarters.

4. Experiment 2

Experiment 2 tested the role of expertise. We wondered whether

counterfactuals focusing on controllable content might be more prevalent when the individual has expertise, i.e., relevant knowledge, skills, and abilities that may provide the basis for subsequent action elaborated within the counterfactual. By contrast, when people find themselves in a situation for which they are ill-equipped for effective action, their counterfactuals might focus less on their own controllable actions and more on the situation or another person's actions. In this logic, one reason why studies by Ferrante et al. (2013, 2016) and Mercier et al. (2016) obtained low rates of controllable counterfactuals might be that the laboratory procedures are experienced by participants as ones for which they lack expertise, whereas under life circumstances for which they are more expert, controllable counterfactuals may prevail.

4.1. Method

The sample was comprised of 99 adult Americans from Mechanical Turk. The initial sample was $N = 102$; 3 participants were excluded for failure to follow instructions (e.g., failing to record an outcome or a counterfactual either by entering gibberish into the response box or leaving it blank). Of the final sample, 53 were women and 46 were men, with age $M = 34.7$. The survey ran Oct 10, 2016.

The outcome probe was similar to that of Experiment 1, and formed the basis of a within-subject manipulation of expertise (order randomized): "Think back to a recent negative experience. This negative experience should involve an event or activity for which you have a lot of expertise and knowledge (i.e., you're an old hand at it) [very little expertise and knowledge (i.e., you're new to it).]" Next, participants responded to the counterfactual probe and dichotomous counterfactual controllability judgment as per the pretest and Experiment 1.

4.2. Results

Expertise had no impact on the rate of controllable counterfactuals. The expert-focused condition evoked controllable counterfactuals at a rate (74%) just as numerous as by the inexpert-focused condition (72%), McNemar $\chi^2 = 0.03$, $p = 0.864$, odds ratio = 1.13. As before, we also analyzed just the first outcome (i.e., expert or inexpert events/activities) that participants responded to. A similar result appeared in this between-subject analysis, such that both expert versus inexpert episodes elicited high rates of controllable counterfactuals (68% vs. 78%), Pearson χ^2 (1, $N = 99$) = 1.33, $p = 0.25$, odds ratio = 0.59.

This experiment indicated that expertise is less viable as an explanation for the discrepancy in rates of controllable counterfactuals. Also, with the same retrospective recall paradigm we again noted that the rate of controllable counterfactuals following a negative, personally meaningful outcome hovers around three quarters.

5. Experiment 3

Experiment 3 tested the role of self-initiation, that is, whether or not the episode in question was one triggered by the individuals' own initiative or was one they were drawn into by others. Compared to other-initiated episodes, self-initiated episodes may enable the individual to draw on their own actions, preferences, and knowledge as input into controllable counterfactual thoughts. Self-initiation might explain the discrepancy between research findings in that the more recent laboratory demonstrations involve tasks that participants did not think to engage in themselves but rather were asked to engage in by experimenters, thus pushing them away from generating internal and controllable counterfactuals. We tested this hypothesis in the same manner as in the previous experiments, via a within-subject manipulation of self- versus other-initiated episode within the context of retrospective self-reports of meaningful life experiences.

5.1. Method

The sample comprised 95 adult Americans from Mechanical Turk. The initial sample was $N = 101$; 6 participants were excluded for failure to follow instructions (e.g., failing to record an outcome or a counterfactual either by entering gibberish into the response box or leaving it blank). Of the final sample, 50 were women and 45 were men, with age $M = 35.7$. The survey ran Oct 25, 2016.

The outcome probe was similar to that of Experiments 1 and 2, providing the basis of a within-subject manipulation of self- versus other-initiated circumstance (order randomized): "Think back to a recent negative experience. This negative experience should be an unusual/peculiar situation that you were drawn into by YOUR OWN [SOMEONE ELSE'S] decision." Next, participants responded to the counterfactual probe and dichotomous counterfactual controllability judgment as per the previous experiments.

5.2. Results

Self-initiated episodes resulted in a greater rate of controllable counterfactuals (86%) than did other-initiated episodes (64%), McNemar $\chi^2 = 10.81$, $p < 0.001$, odds ratio = 3.63. As before, we also analyzed just the first outcome (i.e., self-initiated or other-initiated outcomes) to which participants responded. A similar result was seen in this between-subject analysis, such that there was a higher rate of controllable counterfactuals for self-initiated episodes (85%) than for other-initiated episodes (70%), Pearson χ^2 (1, $N = 95$) = 3.01, $p = 0.08$, odds ratio = 2.38.

A concern that might be raised about the effect of self-initiation on counterfactuals is that of tautology. That is, if under the condition of self-initiation, counterfactuals tend to focus precisely on the self-action that initiated the episode, then we will merely have shown that steering people to focus on X causes them to report back thoughts about X. By contrast, our conception is more general, that self-initiation results in a focus not only on initiating actions, but also actions at any point during the course of an episode. To address this concern about tautology, two independent raters coded the participant's counterfactuals as to whether they focused on the initiation-point itself, some other aspect of the episode, or "don't know/can't say" ($\kappa = 0.69$; 95% CI: 0.53–0.84; 85% agreement). Disagreements were resolved by a third rater. Results showed that counterfactual thoughts based on self-initiated events were more likely to focus on some other aspect of the event (68%), as opposed to the initiation-point itself (28%), with few uncodeable counterfactuals (3%). To examine the impact on controllability, we re-ran the main analyses without the initiation-focused or uncodeable counterfactuals. As before, self-initiated outcomes involved a greater rate of controllable counterfactuals (82%) than did other-initiated outcomes (65%), McNemar $\chi^2 = 4.00$, $p = 0.05$, odds ratio = 2.57.

This experiment therefore reveals evidence that self-initiation plays a role in moderating the extent to which negative episodes evoke controllable counterfactuals, thus accounting in part for the discrepancy in past findings.

6. Experiment 4

The previous experiments employed a dependent measure that prompted participants for counterfactual thoughts. Although such a method has been used widely in previous research to reveal many important effects, it has the drawback of potentially leading participants to a certain way of reporting their thoughts. For example, pushing people to report a counterfactual might implicitly suggest that they focus on their own controllable action. By contrast, an open-ended thought-listing measure would have the advantage of not directing participants to report any particular thought content (see Roese & Olson, 1997 for previous use of such an approach). The drawback of an open-ended measure is that participants tend to report

far fewer counterfactuals spontaneously than when prompted, thus substantially weakening the power to detect effects. Nevertheless, if an open-ended measure revealed essentially the same pattern involving self-initiation observed in Experiment 3, we would have further evidence with which to triangulate upon a conclusion regarding the role of self-initiation. Such was the goal of Experiment 4, which used the same design as Experiment 3 but substituted an open-ended rather than prompted measure of counterfactual thinking.

6.1. Method

The sample comprised 96 adult Americans from Mechanical Turk. The initial sample was $N = 99$; 3 participants were excluded for failure to follow instructions (e.g., failing to record an outcome or thought-listing either by entering gibberish into the response box or leaving it blank). Of the final sample, 48 were women and 48 were men, with age $M = 35.8$. The survey ran April 13–14, 2017.

The outcome probe was similar to that of the previous experiment, providing the basis of a within-subject manipulation of self- versus other-initiated epsiode (order randomized): "Think back to a recent negative experience. Also, this negative experience should involve a situation that you were drawn into by YOUR OWN [SOMEONE ELSE'S] decision or actions." Next, to capture spontaneous counterfactual thoughts, participants responded to an open-ended thought listing task, in which they were given three boxes and asked to provide three thoughts that they had had in response to the event they recorded (one complete thought in each box).

6.2. Results

Two independent raters (blind to experimental hypotheses) coded each thought listed for counterfactual content. Raters were instructed to code a response as a counterfactual only when there was clear evidence that an alternative to reality had been considered ($\kappa = 0.70$; 95% CI: 0.60–0.79; 94% agreement). Each counterfactual was also coded for controllability ($\kappa = 0.34$; 95% CI: 0.23–0.44; 79% agreement), using the same descriptions provided in the dichotomous measure in the previous experiments. Finally, as in Experiment 3, each counterfactual thought about the self-initiated episode was coded as to whether the counterfactual focused on the initiation-point itself, some other facet or time point of the episode, or "don't know/can't say" ($\kappa = 0.71$; 95% CI: 0.52–0.91; 86% agreement). All disagreements were resolved by a third coder.

We first assessed whether the overall rate with which counterfactuals were spontaneously mentioned (i.e., mean number of counterfactuals) varied by the self-initiation manipulation. A repeated-measures ANOVA showed that self-initiated (vs. other-initiated) episodes resulted in more spontaneous counterfactual thoughts (*M*s = 0.51 vs 0.25, *SD*s = 0.70 and 0.56), $F(1, 95) = 9.59$, $p = 0.003$, $\eta_p^2 = 0.09$. When looking only to those counterfactuals deemed to be controllable by the independent coders, we again noted a greater number for self-initiated (vs. other-initiated) episodes (*M*s = 0.39 vs 0.20, *SD*s = 0.62 and 0.52), $F(1, 95) = 6.88$, $p = 0.01$, $\eta_p^2 = 0.07$. By contrast, no significant variation as a function of self-initiation occurred for those counterfactuals deemed to be uncontrollable (*M*s = 0.13 vs 0.05, *SD*s = 0.39 and 0.22), $F(1,95) = 2.37$, $p = 0.13$, $\eta_p^2 = 0.02$.

As in the previous experiments, to check robustness we also analyzed just the first episode (i.e., self-initiated or other-initiated) that participants described. Similar results were seen in this between-subject analysis, such that self-initiated episodes involved more spontaneous counterfactual thoughts overall (*M*s = 0.51 vs 0.16, *SD*s = 0.62 and 0.47), $F(1,95) = 9.57$, $p = 0.003$, $\eta_p^2 = 0.09$, which was driven by a significant difference in spontaneous controllable counterfactuals (*M*s = 0.38 vs 0.10, *SD*s = 0.57 and 0.37), $F(1,95) = 8.24$, $p = 0.005$, $\eta_p^2 = 0.08$, but not spontaneous uncontrollable counterfactuals (*M*s = 0.13 vs 0.06, *SD*s = 0.34 and 0.24), $F(1,95) = 1.24$, $p = 0.27$, $\eta_p^2 = 0.01$.

As in Experiment 3, we tested the issue of tautology in self-initiation effects. As before, results showed that counterfactual thoughts based on self-initiated events were more likely to focus on some other aspect of the event (59%), as opposed to the initiation-point itself (41%), with no uncodeable counterfactuals. To examine the impact on controllability, we re-ran the main analyses without the initiation-focused counterfactuals. The repeated-measures ANOVA showed that this difference was no longer statistically significant, (*M*s = 0.18 and 0.20, *SD*s = 0.44 and 0.52), $F(1, 95) = 0.10$, $p = 0.75$, $\eta_p^2 = 0.001$. We emphasize that with relatively few counterfactual thoughts reported by participants (total counterfactuals per event recall, $M = 0.38$, $SD = 0.64$), the power to detect effects is diminished relative to our other experiments.

7. Experiment 5

Having found evidence to support the role of self-initiation but not normality and expertise as factors that may affect the rate of controllable counterfactuals, the goal of the present experiment was to replicate further the effect of self-initiation on counterfactual content and then to extend that finding by clarifying which attribution dimension(s) display the effect. Recall that the current crop of research has used the construct of personal control, which conflates two of three dimensions of causal explanation specified by Weiner's (1985, 1986) attribution theory, namely locus and controllability. Accordingly, we do not yet know whether self-initiation produces effects on locus, controllability, or stability. By asking three separate questions that tap into each of these attribution dimensions, we may better characterize the effect of self-initiation on the content of counterfactual thinking.

7.1. Method

The sample comprised 99 adult Americans from Mechanical Turk. The initial sample was $N = 100$; 1 participant was excluded for failure to follow instructions (leaving response fields blank). Of the final sample, 49 were women and 50 were men, with age $M = 33.4$. The survey ran December 7, 2016.

The same outcome probe as in previous experiments was used, and the within-subject manipulation of self-initiation (order randomized) was identical to Experiment 3. Next, participants responded to the following questions relating to attribution dimensions. For locus (internal vs. external), the question was: "Does your 'if only' thought focus on something within you (for example, your abilities, skills, or an action you could have taken) as opposed to another person?" For controllability, the question was: "Does your 'if only' thought focus on something that is changeable or controllable by a person (such as effort) or something no person could control (such as game rules or the weather)?" For stability, the question was: "Does your 'if only' thought focus on something that, generally speaking, can change over time and across situations (for example, amount of effort) or something that does rarely changes (for example, height or talent)?".

7.2. Results

Self-initiated outcomes resulted in significant variation on the attribution dimensions of locus and stability, but not controllability. By the far the biggest effect was on locus, such that self-initiated episodes involved counterfactuals that nearly always were internal (93%) whereas other-initiated episodes were less likely to be internal (44%), McNemar $\chi^2 = 44.2$, $p < 0.001$, odds ratio = 49.0. The direction of this effect replicates the finding from the previous two experiments. By contrast, the self-initiation manipulation had no effect on controllability; in both conditions the counterfactuals focused mostly on controllable as opposed to uncontrollable aspects (89% vs. 83%), McNemar $\chi^2 = 1.14$, $p = 0.286$, odds ratio = 1.75. This result unpacks the finding of the previous two experiments to show that variation in the

conflated variable of personal controllability involves variation in the attribution dimension of locus but not controllability. In short, what varies most is whether counterfactuals focus internally versus externally. Self-initiated circumstances influenced the stability dimension, such that self-initiation involved counterfactuals more likely to focus on unstable aspects (91%) than did other-initiated circumstances (77%), McNemar $\chi^2 = 7.68$, $p = 0.006$, odds ratio = 4.5.

To check robustness of these results, as before we also assessed just the first episode that participants described (i.e., self-initiated or other-initiated) in a between-subject analysis. For locus, there was a higher rate of internal-focused counterfactuals for self-initiated outcomes (94%) than for other-initiated outcomes (57%), Pearson χ^2 (1, $N = 99$) = 18.31, $p < 0.0001$, odds ratio = 11.75. For controllability, both self-initiated (90%) and other-initiated (92%) outcomes focused mainly on controllable aspects, Pearson χ^2 (1, $N = 99$) = 0.10, $p = 0.75$, odds ratio = 0.80. For stability, both self-initiated (90%) and other-initiated (86%) outcomes focused mainly on unstable aspects, Pearson χ^2 (1, $N = 99$) = 0.43, $p = 0.51$, odds ratio = 1.5. Summarizing, this between-subject (and hence less powerful) analysis confirmed the results involving the locus and controllability dimensions, but did not confirm the result on the stability dimension.

To test the issue of tautology in self-initiation effects, we again had two independent raters code counterfactual thoughts within the self-initiation condition as to whether the counterfactual focused on the initiation-point itself, some other aspect of the event, or "don't know/can't say" (κ = 0.65; 95% CI: 0.50–0.80; 83% agreement). Disagreements were resolved by a third rater. Results showed that counterfactual thoughts were more likely to focus on some other aspect of the episode (64%) as opposed to the initiation-point itself (35%), with one uncodeable counterfactual (1%). We re-ran the main analysis while excluding the initiation-focused or uncodeable counterfactuals. As before, self-initiation resulted in significant variation on the attribution dimensions of locus and stability, but not controllability. For locus, there was a higher rate of internal-focused counterfactuals (89%) than for other-initiated outcomes (41%), McNemar $\chi^2 = 26.28$, $p < 0.0001$, odds ratio = 31.0. For controllability, both self-initiated (86%) and other-initiated (83%) outcomes focused mainly on controllable aspects, McNemar $\chi^2 = 0.06$, $p = 0.80$, odds ratio = 1.29. For stability, self-initiated outcomes (91%) focused more on unstable aspects than did other-initiated outcomes (76%), McNemar $\chi^2 = 4.27$, $p = 0.039$, odds ratio = 4.0. Taking these results alongside those of Experiments 3 and 4, it seems clear that although self-initiation does prompt a sizable number of counterfactual responses to focus specifically on the initiation of the episode in question, it also pushes a wider focus on one's own actions throughout the episode.

8. General discussion

Motivated by the functional theory of counterfactual thinking (Epstude & Roese, 2008, 2011; Roese & Epstude, 2017), the present research explored determinants of the extent to which episodic counterfactual thinking focuses on internally controllable action. According to the functional theory, episodic counterfactual thoughts are generated in the service of goal pursuit, the most common of which centers on preparation for personal improvement. Thus, counterfactual thoughts typically focus on a particular means for achieving an end, and the process of generating counterfactual thoughts contributes, under some conditions, to performance improvement and goal success. Recent research has challenged a core tenet of the functional theory of counterfactual thinking: that episodic counterfactual thoughts may influence performance by way of the content-specific pathway (i.e., that counterfactuals contain a causal insight specifying how a personal action can bring about a desired outcome). That is, for episodic counterfactuals to fulfill a preparatory function via this content-specific pathway, they must focus on actions that are under the control of the individual. The content-neutral pathway by which counterfactual thinking can influence performance, which involves generic motivational effects, has not been challenged by this new research and was not the focus of the present research.

The theoretical puzzle is that a range of earlier findings indicated that episodic counterfactual thoughts largely focus on internally controllable actions (Davis et al., 1995; Girotto et al., 1991; Hammell & Chan, 2016; Mandel, 2003; Mandel & Lehman, 1996; Markman et al., 1995; McCloy & Byrne, 2000; McEleney & Byrne, 2006; Morris et al., 1999; Roese & Olson, 1995), whereas newer research challenges this conclusion in showing that episodic counterfactuals largely focus on external factors (Ferrante et al., 2013, 2016; Girotto et al., 2007; Mercier et al., 2016; Pighin et al., 2011). Another part of the puzzle is how best to explain the considerable variability in the proportion of controllable counterfactuals across the latter collection of studies (ranging from 9% in Mercier et al., 2016, Study 1b to 68% in Ferrante et al., 2016, Study 2; see Table 1 for summary). We reiterate our clarification that some of this past research conflates two separable causal dimensions, that of locus (internal vs. external) and controllability (controllable vs. uncontrollable). The present research offers a solution to this puzzle in terms of the role of self-initiation. The more that an episode is self-initiated, the more likely it is that resultant counterfactual thinking will center on internal rather than external factors. In our experiments, episodic counterfactuals tended overwhelmingly to focus on controllable factors irrespective of self-initiation.

The present research offered three noteworthy findings. Using a methodology in which participants recall meaningful events from autobiographical memory, a wide variety of episodic counterfactual thoughts were reported by participants. The first noteworthy finding was that across all experiments we found a strong, general tendency for episodic counterfactuals to focus on internal and controllable action (often at a proportion of about three fourths). We argue that this finding contains considerable generality, because the retrospective self-report methodology permitted participants to describe events from a wide range of personally experienced life events. These are the sorts of circumstances that evoke counterfactual thinking that connects to personal goals and hence drive a focus on internal and controllable action.

The second noteworthy finding was the systematic test of three potential determinants – normality, expertise, and self-initiation – that may plausibly explain variation across past reported results. We manipulated these three factors in successive experiments to specify which best accounted for variability in the rate at which counterfactuals focus on internal and controllable action. We found no evidence for the determinant roles of normality and expertise, but we did find evidence for the role of self-initiation. Experiments 3, 4, and 5 showed that self-initiated episodes involve a greater focus of counterfactuals on internal (vs. external) than do other-initiated episodes. When circumstances are engineered by the individual, as when he or she decides to go to a movie, or shops for gifts for loved ones, or selects the groceries for the week's cooking, problems that then arise evoke counterfactual thoughts that center specifically on what he or she could do differently to have achieved a more desirable state of affairs. By contrast, when circumstances are engineered by something or someone else, as when taking an exam written by an instructor, going to movie selected by a friend, or dining at a sibling's residence, counterfactuals may focus more on facets not centering on the individual.

An important methodological clarification addressed the possibility that self-initiation effects were merely a tautological demonstration. That is, if after a self-initiated episode, counterfactuals focus on the self-action that initiated the episode, then we may merely have shown that steering people to focus on X causes them to report thoughts about X. In content coding of counterfactuals in Experiments 3, 4, and 5, we found that although self-initiation does prompt a sizable number of counterfactual responses to focus on the initiation of the episode in question, it also pushes a wider a focus on one's own actions throughout the episode. In applying this observation to past studies, it is plausible that

counterfactual thoughts about the episode initiation may have been captured in previous studies and could therefore partially explain the discrepancies in rates of internally controllable counterfactual responses.

The present findings thus explain at least some of the puzzle that initiated this research, namely that a recent crop of experiments showed that episodic counterfactual tended to focus on aspects outside of the individual's control (see Table 1). Looking across past research, those that involved situations high in self-initiation do seem to show a greater emphasis on counterfactuals describing internally controllable action (Davis et al., 1995, Study 2; Girotto et al., 1991; Mandel, 2003; Mandel & Lehman, 1996; McCloy & Byrne, 2000; McEleney & Byrne, 2006; Roese & Olson, 1995), whereas those that involved situations low in perceived self-initiation show a weaker emphasis on counterfactuals describing internally controllable action (Epstude & Jonas, 2015; Ferrante et al., 2013; Ferrante et al., 2016; Mercier et al., 2016; Pighin et al., 2011). To be sure, there are exceptions and ambiguities in looking across past research. For example, the gambling task used by Markman et al. (1995) was likely low in self-initiation because participants did not select this task in particular, yet their counterfactuals nevertheless focused on aspects of the procedure under their direct control. We do not claim that self-initiation explains all variation in past experimental results, only some of it.

The third noteworthy finding was the clarification of what is meant by "personally controllable action" by unpacking this construct into the three dimensions of causal attribution specified in Weiner's (1985, 1986) theory: i.e., locus, controllability, and stability. According to this theory, the dimension of locus (differentiating a focus internal versus external to oneself) is orthogonal to controllability (an aspect may be controllable by a person vs. uncontrollable by a person), and stability (differentiating how much or how often the aspect changes, as captured in the difference between effort and ability). In some of the counterfactual publications we have reviewed, the dimensions of locus and controllability have been conflated. The theoretical question, then, is whether we may be seeing variation across episodes and studies in terms of locus, or controllability, or both. Experiment 5 provided this disambiguation by assessing variation on all three of Weiner's dimensions as a function of self-initiation, the same determinant that had explained significant variation in Experiments 3 and 4. Thus, Experiment 5 replicated the results of Experiments 3 and 4, but in addition added further clarity by showing that variation occurred on the locus and stability dimensions but not on the controllability dimension. As circumstances became less self-initiated, the proportion of counterfactuals focusing on internal aspects dropped. At the same time, those counterfactuals remained just as firmly fixated on controllable action; however, those controllable actions just happened to be under someone else's control.

Accordingly, an important contribution of the present work is to clarify that episodic counterfactuals focus on controllable (vs. uncontrollable action) most of the time and that this effect is stubbornly resistant to attempts to manipulate it. Further, episodic counterfactuals do also tend to focus on one's own personal actions, but as the circumstances involve less self-initiation, the focus of counterfactuals reorients from controllable actions of the self to the controllable actions of others. These results suggest the possibility that a great many prior findings discussed in terms of variation in controllability in fact involved variation in locus of causation. Such a possibility represents an extrapolation of variation stemming from the factor of self-initiation to other plausible factors that may influence counterfactual thinking, and thus remains tentative pending direct tests involving those other factors. Accordingly, we recommend that future research ensure the specific and separable measurement of episodic counterfactual content along all three attribution dimensions of locus, controllability, and stability.

A further contribution was to connect the attributional dimension of stability to counterfactual content. Although part of Weiner's (1985, 1986) theory of causal attribution, the stability construct per se has not been explored in the counterfactual literature. However, as we previously noted, the widely reported finding that counterfactual content tends to seize upon abnormal (more than normal) antecedent factors (e.g., Gavanski & Wells, 1989; Kahneman & Tversky, 1982; Miller & McFarland, 1986; Roese & Hur, 1997) may be interpreted as a more general observation that counterfactuals tend to focus on that which is unstable or dynamic (rather than fixed, normal, or typical). Here we show using a rather more direct approach that counterfactual thoughts generally focus on unstable causes, i.e., aspects that change and vary over time (e.g., effort, rather than ability). It bears stating that from a theoretical standpoint, counterfactuals and causal attributions are closely related, in that episodic counterfactuals often embody a causal inference connecting an action to a desired outcome (e.g., "If I had done X, it would have brought about Y"). It therefore behooves counterfactual theorists to attend to the causal attribution literature to forge new theoretical integrations. Weiner (1986) summarized earlier studies of spontaneous causal inferences as centering mostly on internal rather than external factors, and also on unstable rather than stable factors. More recently, Malle (2006) (see also Malle, 2011) concluded that "actors, compared with observers, prefer to explain their intentional actions with reference to the reasons for which they acted – the mode of explanation that emphasizes the agent's control, deliberation, and free choice" and "estimated from about 10,000 verbal explanations I have collected over the past 10 years, only 5%-10% of all explanations refer to personality traits, and an additional 5% refer to stable beliefs and preferences" (p. 911). The present results, involving episodic counterfactuals generated with respect to autobiographical memory of experienced events, are largely similar in their emphasis on internal, controllable, and unstable dimensions of causal explanation.

This research clarifies the functional theory of counterfactual thinking (Epstude & Roese, 2008, 2011; Roese, 1994, 1997, 1999; Roese & Epstude, 2017), which seeks to describe observable patterns in terms of goal cognition. According to the theory, counterfactual thoughts may bring about performance improvement via either a content-specific pathway (in which the counterfactual insight directly informs behavior change by specifying the particular means) or a content-neutral pathway (in which activation by counterfactual thought of cognitive procedures or negative affect bring about behavior change independently of the specific informational content of the counterfactual itself). The present findings clarify the content-specific pathway in which counterfactuals that focus on internally controllable action constitute a form that is better suited to serving the function of goal progress because insights that are specific to the individual and specific to controllable action are most easily applied to new plans and intentions. Moreover, self-initiation heightens a focus on individual action because, we argue, self-initiation brings with it a palette of familiar behavioral options upon which counterfactuals may focus.

We had initially posited that normality and expertise might also feed into the tendency to focus on internally controllable counterfactuals, but our experiments indicated that these factors account for far less variation in counterfactual content than does self-initiation. More broadly, the functional theory suggests that the preparative (versus affect-regulatory or self-presentational functions) is the default for counterfactual thinking, in part because the upward (vs. downward) counterfactual is the most prevalent form. The present finding, that recalled episodic counterfactuals for the most part focus on factors that are internal, controllable, and unstable adds further, convergent support for the contention that the preparative function is the default for counterfactual thinking. At the most general level, episodic counterfactuals connect to perceptions of human agency and the possibilities open to people who dream of better things (Alquist, Ainsworth, Baumeister, Daly, & Stillman, 2015; Kray et al., 2010; Kulakova, Khalighinejad, & Haggard, 2017; Waytz, Hershfield, & Tamir, 2015).

To summarize, the present research was guided by the functional theory of counterfactual thinking and was executed by way of

tabulating episodic counterfactual thoughts generated in relation to autobiographical memories. We found that counterfactuals focus on internal and controllable actions the majority of the time. Explaining the discrepancy in past findings, we found that counterfactuals are more likely to focus on internal (vs. external) aspects to the extent that the episode was self-initiated.

References

Alquist, J. L., Ainsworth, S. E., Baumeister, R. F., Daly, M., & Stillman, T. F. (2015). The making of might-have-beens. *Personality and Social Psychology Bulletin*, *41*, 268–283.

Bear, A., & Knobe, J. (2016). Normality: Part descriptive, part prescriptive. *Cognition*. http://dx.doi.org/10.1016/j.cognition.2016.10.024.

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157.

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.

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.

Epstude, K., & Jonas, K. J. (2015). Regret and counterfactual thinking in the face of inevitability. *Social Psychological and Personality Science*, *6*, 157–163.

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

Epstude, K., & Roese, N. J. (2011). When goal pursuit fails: The functions of counterfactual thought in intention formation. *Social Psychology*, *42*, 19–27.

Epstude, K., Scholl, A., & Roese, N. J. (2016). Prefactual thoughts: Mental simulations about what might happen. *Review of General Psychology*, *20*, 48–56.

Ferrante, D., Girotto, V., Stragà, M., & Walsh, C. R. (2013). Improving the past and the future: A temporal asymmetry in hypothetical thinking. *Journal of Experimental Psychology: General*, *142*, 23–27.

Ferrante, D., Stragà, M., Walsh, C., & Girotto, V. (2016). *What could I have done or what can I do? The effect of counterfactual and prefactual thinking on predictions and intentions. Paper presented at the EASP Small Group Meeting "Counterfactual thinking in causality, emotion, communication, and behavior", Aix-en-Provence, France.* (2016, June).

Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, *25*, 314–325.

Girotto, V., Ferrante, D., Pighin, S., & Gonzalez, M. (2007). Postdecisional counterfactual thinking by actors and readers. *Psychological Science*, *18*, 510–515.

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, *78*, 111–133.

Hafner, R. J., White, M. P., & Handley, S. J. (2012). Spoilt for choice: The role of counterfactual thinking in the excess choice and reversibility paradoxes. *Journal of Experimental Social Psychology*, *48*, 28–36.

Hammell, C., & Chan, A. Y. C. (2016). Improving physical task performance with counterfactual and prefactual thinking. *PloS One*, *11*(12), e0168181.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.

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

Kray, L. J., George, L. G., Liljenquist, K. A., Galinsky, A. D., Tetlock, P. E., & Roese, N. J. (2010). From what might have been to what must have been: Counterfactual thinking creates meaning. *Journal of Personality and Social Psychology*, *98*, 106–118.

Kulakova, E., Khalighinejad, N., & Haggard, P. (2017). I could have done otherwise: Availability of counterfactual comparisons informs the sense of agency. *Consciousness and Cognition*, *49*, 237–244.

Malle, B. F. (2006). The actor-observer asymmetry in causal attribution: A (surprising) meta-analysis. *Psychological Bulletin*, *132*, 895–919.

Malle, B. F. (2011). Time to give up the dogmas of attribution: An alternative theory of behavior explanation. *Advances of Experimental Social Psychology*, *44*, 297–352.

Mandel, D. R. (2003). Counterfactuals, emotions, and context. *Cognition and Emotion*, *17*, 139–159.

Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology*, *71*, 450–463.

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.

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1995). The impact of perceived control on the imagination of better and worse possible worlds. *Personality and Social Psychology Bulletin*, *21*, 588–595.

McCloy, R., & Byrne, R. M. J. (2000). Counterfactual thinking about controllable events. *Memory and Cognition*, *28*, 1071–1078.

McEleney, A., & Byrne, R. M. J. (2006). Spontaneous causal and counterfactual thoughts. *Thinking and Reasoning*, *12*, 235–255.

Mercier, H., Rolison, J. J., Stragà, M., Ferrante, D., Walsh, C. R., & Girotto, V. (2016). Questioning the preparatory function of counterfactual thinking. *Memory and Cognition*, *45*, 261–269.

Miller, D. T., & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin*, *12*, 513–519.

Morris, M. W., Moore, P. C., & Sim, D. L. H. (1999). Choosing remedies after accidents: Counterfactual thoughts and the focus on fixing "human error". *Psychonomic Bulletin and Review*, *6*, 579–585.

Özbek, M., Bohn, A., & Berntsen, D. (2016). Imagining the personal past: Episodic counterfactuals compared to episodic memories and episodic future projections. *Memory and Cognition*.

Petrocelli, J. V., & Harris, A. K. (2011). Learning inhibition in the Monty Hall problem: The role of dysfunctional counterfactual prescriptions. *Personality and Social Psychology Bulletin*, *37*, 1297–1311.

Pighin, S., Byrne, R. M., Ferrante, D., Gonzalez, M., & Girotto, V. (2011). Counterfactual thoughts by experienced, observed and narrated events. *Thinking and Reasoning*, *17*, 197–211.

Quelhas, A. C., & Byrne, R. M. J. (2003). Reasoning with deontic and counterfactual conditionals. *Thinking and Reasoning*, *9*, 43–65.

Revlin, R., Cate, C. L., & Rouss, T. S. (2001). Reasoning counterfactually: Combining and rending. *Memory and Cognition*, *29*, 1196–1208.

Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, *66*, 805–818.

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*, 133–148.

Roese, N. J. (1999). Counterfactual thinking and decision making. *Psychonomic Bulletin and Review*, *6*, 570–578.

Roese, N. J., & Epstude, K. (2017). The functional theory of counterfactual thinking: New evidence, new challenges, new insights. *Advances in Experimental Social Psychology*.

Roese, N. J., & Hur, T. (1997). Affective determinants of counterfactual thinking. *Social Cognition*, *15*, 274–290.

Roese, N. J., & Olson, J. M. (1995). Outcome controllability and counterfactual thinking. *Personality and Social Psychology Bulletin*, *21*, 620–628.

Roese, N. J., & Olson, J. M. (1997). Counterfactual thinking: The intersection of affect and function. *Advances in Experimental Social Psychology*, *29*, 1–59.

Rye, M. S., Cahoon, M. B., Ali, R. S., & Daftary, T. (2008). Development and validation of the counterfactual thinking for negative events scale. *Journal of Personality Assessment*, *90*, 261–269.

Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of Learning and Memory*, *117*, 14–21.

Summerville, A. (2011). The rush of regret: A longitudinal analysis of naturalistic regrets. *Social Psychological and Personality Science*, *2*, 627–634.

Thompson, V., & Byrne, R. M. J. (2002). Reasoning counterfactually: Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1154–1170.

Waytz, A., Hershfield, H. E., & Tamir, D. I. (2015). Mental simulation and meaning in life. *Journal of Personality and Social Psychology*, *108*, 336–355.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548–573.

Weiner, B. (1986). *An attributional theory of motivation and achievement*. New York: Springer-Verlag.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. *Advances in Experimental Social Psychology*, *35*, 345–411.