Counterfactual Thinking and Self-Motives

Lawrence J. Sanna  
*Institute for Social Research, University of Michigan*

Edward C. Chang  
*University of Michigan*

Susanne Meier  
*Washington State University*

Three studies indicated that valenced events and self-motives have implications for understanding processes underlying counterfactual thinking. Moods (Study 1) and outcome valence (Study 3) influenced counterfactuals when self-motives (self-improvement, mood-repair, mood-maintenance, and self-protection) were manipulated directly. Agreement and reaction times (Studies 1 and 2), as well as time pressure (Study 3), indicated that counterfactual responses can be quick or slow depending on whether self-motives suggest a direction either consistent or inconsistent with direction activated initially. In Study 2, responses to manipulated outcomes by high- and low-self-esteem persons, who differ naturally in self-motives, provided further evidence for proposals when task repeatability was varied. Implications for antecedents and consequences of counterfactual thinking, self-motives, and dual-process models, are discussed.

Our lives are filled with events that make us feel bad or good. A bungled employment interview, a fortuitous chance encounter, a horribly failed exam, or a lovers’ compromise are but a few examples of common occurrences that can influence our feelings. However, not only is it what actually transpired that matters; people are influenced by counterfactual mentally simulated alternatives that did not really happen but that easily can be imagined having happened instead. These thoughts about “what might have been” can occur spontaneously (Sanna & Turley, 1996) and are often characterized by “if only,” “at least,” or similar conceptions (see Miller, Turnbull, & McFarland, 1990; Roese, 1997, for reviews). We here propose and test an integrative framework that distinguishes counterfactuals on the basis of whether they are activated by situational (e.g., outcomes or moods) or characterological (e.g., self-motives or personality) attributes. Not only might this synthesis extend knowledge of counterfactual antecedents and consequences but it also might broaden an understanding of how counterfactuals function to serve diverse self-motives.

Valenced Events and Counterfactual Direction

Valenced events are unavoidable. Assessing people’s responses to them can contribute to an understanding of coping and well-being. Counterfactual researchers have varied valenced events by either outcomes or moods. Markman, Gavanski, Sherman, and McMullen (1993) tested how outcomes influence counterfactuals by having participants play a computer-simulated blackjack game. Upward counterfactuals are simulated alternatives better than actuality (e.g., “If only I made that last free-throw, we would have won the game”); downward counterfactuals are simulated alternatives worse than actuality (e.g., “At least my smoke detector worked, or I might have been killed”). Outcomes were framed as a failure (loss) or success (win), and participants thought they would or would not have a second try. Failures resulted in upward counterfactuals when participants had a second try, and successes resulted in downward counterfactuals. Other research has replicated this.

Authors’ Note: This research was completed while Lawrence J. Sanna was on leave from the Department of Psychology, Washington State University. We thank anonymous reviewers for their comments on a previous version of this article and the following people for their assistance with data collection and coding: Melissa Gilmore, Mark Hutchinson, Joshua Lechter, Christina Tortora, and Eric Wegner. Correspondence concerning this article should be addressed to Lawrence J. Sanna, who is now at the Department of Psychology, CB #3270 Davie Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3270; e-mail: sanna@unc.edu; Web site: http://www.unc.edu/~sanna.

*PSPB, Vol. 27 No. 8, August 2001* 1023-1034  © 2001 by the Society for Personality and Social Psychology, Inc.
pattern of upward counterfactuals after failures and downward counterfactuals after successes (e.g., Roese & Olson, 1995; Sanna, 1996; Sanna & Turley-Ames, 2000).

Moods have produced an identical pattern. Sanna, Turley-Ames, and Meier (1999; Sanna, Meier, & Turley-Ames, 1998) varied moods by films or music. Bad moods induced upward counterfactuals, and good moods induced downward counterfactuals. Moods also influence prefactuals, thoughts of “what may be,” in a similar way (Sanna, 1998, 1999). Connections between outcomes and moods are conspicuous, because numerous life events, like failures and successes, can exert influences through moods (e.g., Brown & Mankowski, 1993). Moods may influence counterfactuals because they inform people’s current states (Schwarz & Clore, 1996), and moods do not have such influences when they are attributed to irrelevant external sources (Sanna et al., 1998). People interpret their lives negatively in bad moods (e.g., “I am a failure” or “There is a problem”) but positively in good moods (e.g., “I am a success” or “Things are fine”), resulting in a pattern similar to outcome valence. Failures or bad moods and successes or good moods leading to upward and downward simulations, respectively, are represented by the gray arrows in Figure 1.

COUNTERFACTUAL DIRECTION AND SELF-MOTIVES

But why is this pattern found? Preparative or self-improvement motives are served by upward counterfactuals, and affective or self-enhancement motives are served by downward counterfactuals (e.g., Markman et al., 1993; Roese, 1994; Sanna, 1996; see also Taylor & Schneider, 1989). Thinking about better things can be a first step to self-improvement. If bad moods signal trouble (Schwarz & Clore, 1996), then upward counterfactuals also can suggest routes for alleviating problems. That Markman et al. found more upward counterfactuals after failures, especially on repeatable tasks, is consistent with this idea. The self-enhancement motive examined has been one of mood-repair (Sanna et al., 1999); after failures or bad moods, downward counterfactuals can be used to feel better. We argue that there is more to counterfactuals than just self-improvement and mood-repair, and hypotheses about diverse self-motives are presented in Table 1. Although there may be several ways to parse self-motives, Sedikides and Strube (1997) also delineated three varieties of self-enhancement: People can repair, maintain, or protect a positive self-concept.

Additional motives may be served by counterfactuals, suggesting novel uses not previously advanced in prior research. At present, there is no explanation for why downward counterfactuals are thought of more after successes and good moods. However, other research indicates that people sometimes wish to maintain or prolong pleasant affective states (e.g., Isen, 1987), a mood-maintenance motive. We argue that mood-maintenance may be achieved by thinking downward counterfactually after positive events or moods. Indirect evidence for this comes from the finding that people enjoy thinking about downward counterfactuals in good moods (Sanna, Meier, & Wegner, in press), perhaps suggestive of mood-maintenance. Self-enhancement may be achieved further by protecting the self from threats, or self-protection. Tests of self-protective uses of mental simulations are scant. However, indirect evidence indicates people may “buffer” or “brace” for failure by lowering confidence and using upward simulations (Sanna, Meier, & Weger, in press), perhaps suggestive of self-protection. We provide direct tests of each of these self-motives.

THE PRESENT RESEARCH

Combining effects of valenced events with predictions about self-motives, our research had several objectives. To summarize, our hypotheses can be viewed in two parts. As described, negative events or moods lead to upward counterfactuals, and positive events or moods lead to downward counterfactuals, represented by the gray arrows of Figure 1. Also as described, diverse self-motives may be served by counterfactuals (i.e., self-improvement, mood-repair, mood-maintenance, self-protection), which may vary by direction as outlined in Table 1. We predicted that if self-motives suggest mental simulations of a direction inconsistent with a direction activated initially, then greater effort is garnered (e.g., counteracting upward with downward simulations to mood-repair), and resulting simulations will be slower. However, if self-motives suggest simulations of a direction consistent with a direction activated initially, then
little additional effort is needed (e.g., mood-maintaining in good moods), and resulting simulations will be quicker. Whether ensuing simulations are quick or slow is represented by the black arrow in Figure 1.

Evidence bearing on our hypotheses is incomplete. Sanna et al.’s (1999) participants agreed to upward counterfactuals fastest in bad moods and downward counterfactuals fastest in good moods. Roese and Hur’s (1997) participants agreed with the statement, “My anagram score could easily have been different,” quickest after failure than success, arguing that faster reactions suggest automatic activation, but no assessment of direction was made. If upward counterfactuals function for self-improvement, they may be brought to mind easily after bad outcomes or moods. Although downward counterfactuals are generated after success, and are agreed with faster in good moods, whether they service mood-maintenance had not been tested previously. Sanna et al. found that while upward counterfactuals were agreed to quickly in bad moods, when mood-repair was salient, upward simulations were overridden with downward ones, with slower responses. But they did not test other self-motives. The present research thus coalesces and moves beyond prior work by proposing and testing a unified framework based on a consistency versus inconsistency idea.

STUDY 1: MANIPULATED SELF-MOTIVES AND MOODS

We used a direct approach in our first study by asking participants specifically to consider self-motives. Hypotheses about links between self-motives and simulations are presented in Table 1. Study 1 used a 4 (self-motive: self-improvement, mood-repair, mood-maintenance, self-protection) × 2 (mood: bad, good) × 2 (counterfactual: upward, downward) mixed design, with counterfactual within-subjects. We predicted fastest reaction times to upward and downward counterfactual statements when self-motives suggest a direction consistent with that activated initially. In bad moods, quick responses should occur for self-improvement and self-protection, but slow responses should occur for mood-repair. In good moods, quick responses should occur for mood-maintenance.

Method

PARTICIPANTS

Participants were 120 (60 women and 60 men) introductory psychology students who received extra course credit. There were equal numbers of participants, and approximately equal numbers of women and men, within each experimental condition.

PROCEDURE

Participants were tested individually and seated at a table with a personal computer. A cover story noted how the researchers were interested in people’s reactions to life events and that test-taking abilities and intellectual performances were of focus. Participants read the following instructions presented by computer (e.g., Sanna & Mark, 1995; Sanna & Turley-Ames, 2000):

We are studying people’s test-taking competence and aptitude on a test of intellectual ability called the Remote Associates Test (RAT). Each RAT item consists of three stimulus words that are somehow related to a fourth word that you are to determine and record. For example, an item might consist of the three stimulus words: elephant, lapse, and vivid. A correct response would be the fourth word memory. That is, in this example, the fourth word, memory, can be related to each of the three stimulus words in the following way: (a) memory like an “elephant,” (b) memory “lapse,” and (c) “vivid” memory. During this experiment, you will be asked to perform a series of RAT items and to answer some questions concerning your perceptions of the tasks and your performance.
RAT. The RAT consists of three separate lists (see McFarlin & Blascovich, 1984). In Study 1, we used the control list, composed of five easy and five difficult items and overall is moderately difficult. Each RAT item consisted of three stimulus words, related to a fourth unreported word that participants identified and recorded. Stimulus words were presented by computer for 1 minute. During each minute, participants attempted to identify the fourth word. If identified, participants typed the word into the computer. If participants could not think of a response, they could leave their answer blank or take a guess; however, instructions stated that each word triad would remain on the screen for only 1 minute. The 10 RAT items were presented to participants in random order. Similar procedures have been used in prior research (Sanna & Mark, 1995; Sanna & Turley-Ames, 2000).1

Mood. As a purported unrelated task, participants’ moods were varied by having them watch and rate films. In the bad-mood condition, participants watched sad clips from the films Gallipoli and Sophie’s Choice, whereas in the good-mood condition, participants watched humorous clips from Splash and Stripes. Preceding these, participants watched a car-chase scene from the movie Bullit; although engaging, this clip is neutral in valence. The series of films lasted about 20 minutes. After each clip, participants responded to “Pilot Movie Ratings,” which asked for routine ratings of the films (e.g., whether they had seen the movie before). These procedures have been used in prior research (Martin, Ward, Achee, & Wyer, 1993; Sanna et al., 1999). Each adjective was rated on 9-point scales anchored by 1 (not at all) and 9 (very much).

Following Martin et al. (1993), participants were then told about another task that was (ostensibly) being pilot tested. Participants drew a map of their university campus for 1 minute. The actual purpose of this task was to create a brief interval between mood ratings and the task of main interest (described next), because a few studies suggest that participants might discount their moods as a basis for behaviors if the moods are rated immediately before proceeding to the task of interest.

Self-motives and counterfactuals. Participants responded to a series of upward or downward counterfactual statements, which were presented by computer (see Sanna & Turley-Ames, 2000; Sanna et al., 1999). To accomplish this, participants read the following:

As part of our study about people’s reactions to various life events, we will provide you with a series of statements about your RAT performance. These statements represent thoughts that some people might have in reaction to their RAT performance. People can have these thoughts for various reasons. We would like you to think back on your RAT performance and respond to each statement by pressing the appropriate numerical keys on your computer keyboard.

Four self-motives were varied. In the self-improvement condition, participants read the following:

As you consider each statement, ask yourself whether such a thought would help you to plan or prepare for the future. That is, think about whether having such thoughts would help you to improve.

In the mood-repair condition, participants read the following:

As you consider each statement, ask yourself whether such a thought would help you to feel better. That is, think about whether having such thoughts would help you to improve your mood.

In the mood-maintenance condition, participants read the following:

As you consider each statement, ask yourself whether such a thought would help keep your mood at its current level. That is, think about whether having such thoughts would help to maintain your mood.

In the self-protection condition, participants read the following:

As you consider each statement, ask yourself whether such thoughts would help you feel better in case you did poorly in the future. That is, consider whether having such thoughts would help to protect your mood.

The remaining instructions were adapted from Schimmack and Diener (1997; see also Sanna & Turley-Ames, 2000) and continued as follows:

Use the numbers from 0 to 9 on the computer keyboard for your responses. A 0 response means that the statement does not characterize a thought that you are having. Responses from 1 to 9 mean that you agree with the statement to the following degrees (1 = very weakly, 5 = moderately, 9 = very strongly); you can use the remaining numbers to indicate more specific degrees of agreement.

A scale ranging from 0 to 9 appeared beneath each statement. Responses were made by pressing the numerical keys on the computer. Each participant was pre-
presented with a set of 10 (5 upward and 5 downward) counterfactuals, the order of which was random, in a within-subjects fashion. Five statements were about better performances (e.g., “If only I had gotten some easier RAT items, I might have performed a lot better”), and five statements were about worse performances (e.g., “I might have performed worse on the RAT if only I had less time”). Statements were randomly selected by computer for each participant from a larger pool of parallel upward and downward statements used in prior research (Sanna & Turley-Ames, 2000; Sanna et al., 1999), with the constraint that no two parallel statements could be used for the same participant. Participants had as much time as they liked to complete the statement-rating task, but they were told to make sure responses reflected thoughts about their RAT performance. Response times and numerical ratings for each statement were recorded.

Results and Discussion

MANIPULATION CHECK

Ratings of negative mood adjectives were reverse scored and averaged with positive mood adjectives (Cronbach’s α = .91). A 4 (self-motive) × 2 (mood) ANOVA revealed a mood main effect, F(1, 112) = 10.82, p < .01. Participants viewing the humorous films (M = 6.22) felt better than those viewing the sad films (M = 3.64), indicating that our mood manipulations were effective.

AGREEMENT AND LATENCY

Important for assessing our hypotheses are specific contrasts (Rosenthal & Rosnow, 1985) of counterfactual direction within cells and between moods (see Table 2). However, separate 4 (self-motive) × 2 (mood) × 2 (counterfactual) ANOVAs on agreement and latency were first conducted to assess other possible effects. Each participant responded to five upward and five downward statements. To account for this, we used average agreement for the set of upward and downward statements in our analyses. We also used averaged reaction times for upward and downward statements in our analyses of latency data.

The ANOVA on agreement revealed Self-Motive × Counterfactual, F(3, 112) = 12.71, p < .01; Mood × Counterfactual, F(1, 112) = 4.33, p < .05; and three-way, F(3, 112) = 7.27, p < .01, interactions. The ANOVA on latency revealed a Mood × Counterfactual interaction, F(1, 112) = 18.58, p < .01. Agreement was greater for upward than downward statements for self-improvement (M_{upward} = 7.83, M_{downward} = 5.08) and self-protection (M_{upward} = 7.32, M_{downward} = 5.61); agreement was greater for downward than upward statements for mood-repair (M_{downward} = 7.08, M_{upward} = 4.41) and mood-maintenance (M_{downward} = 6.55, M_{upward} = 5.21) (comparisons within motives for the first three, ps < .05; comparisons within motives for mood-maintenance, p < .08).

Agreement also was greater for upward (M = 6.84) than downward (M = 5.95) statements in bad moods (p < .05) and downward (M = 6.21) than upward (M = 5.90) statements in good moods (p < .07). Mirroring this pattern, latencies were faster for upward (M = 5.53 s) than downward (M = 7.60 s) statements in bad moods but faster for downward (M = 5.11 s) than upward (M = 7.70 s) statements in good moods (both within-mood ps < .05).

Contrasts within cells and between moods are particularly relevant to hypotheses. As predicted for self-improvement in bad moods, upward counterfactuals were agreed with more than downward ones, and they were agreed with quickly (agreement and latency upward vs. downward ps < .05). Also for self-improvement, although there was a similar agreement pattern in good moods, upward counterfactuals were agreed with slowly (agreement and latency upward vs. downward ps < .05). Between moods, for agreement, the two upward and two downward means did not differ but upward and downward latency means differed (p < .05).

As predicted for mood-repair in bad moods, downward counterfactuals were agreed with more than upward ones, and they were agreed with slowly (agreement and latency downward vs. upward ps < .05). Also for mood-repair was a similar pattern of agreement in good moods, but upward counterfactuals were agreed with slowly (agreement and latency upward vs. downward ps < .05). Between moods for agreement, the two downward means differed (p < .05) but the two upward ones did not. Between moods for latency, the two downward means differed (p < .05) as did the two upward ones marginally (p < .08).

<table>
<thead>
<tr>
<th>TABLE 2: Mean Agreement and Response Latencies for Counterfactual Statements by Self-Motive and Mood for Study 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-Motive</strong></td>
</tr>
<tr>
<td><strong>Agreement</strong></td>
</tr>
<tr>
<td>Self-improvement</td>
</tr>
<tr>
<td>Mood-repair</td>
</tr>
<tr>
<td>Mood-maintenance</td>
</tr>
<tr>
<td>Self-protection</td>
</tr>
<tr>
<td>Agreement</td>
</tr>
<tr>
<td>Latency(s)</td>
</tr>
<tr>
<td>Latency(s)</td>
</tr>
</tbody>
</table>
As predicted for mood-maintenance in good moods, downward counterfactuals were agreed with more than upward ones, and they were agreed with quickly (agreement and latency downward vs. upward ps < .05). Also for mood-maintenance, although not specifically predicted, in bad moods, upward counterfactuals were agreed with more than downward ones, and they were agreed with quickly (agreement and latency upward vs. downward ps < .05). This latter effect mirrors mood-maintenance in good moods. Although not explicitly hypothesized, it is harmonious with our proposals that if counterfactuals activated from moods and self-motives suggest similar directions, they will be agreed with quickly. Mood-maintenance in bad moods means “staying in a bad mood,” and links between upward counterfactuals and bad moods allow this. Between moods, downward and upward agreement and latencies differed (ps < .05). As predicted for self-protection in bad moods, upward counterfactuals were agreed with more than downward ones, and they were agreed with more quickly (agreement and latency upward vs. downward ps < .05). Also for self-protection, although there was a similar pattern of agreement in good moods, upward counterfactuals were agreed with more slowly (agreement and latency upward vs. downward ps < .05). Between moods, for agreement, the two upward and two downward means did not differ, but latency means differed (ps < .05).

Self-motives and moods influenced counterfactuals in Study 1, with several findings of note. First, ANOVAs supported hypotheses in Table 1 about relations between self-motives and counterfactual direction. Second, ANOVAs revealed that upward counterfactuals were agreed with more and faster in bad moods but downward counterfactuals were agreed with more and faster in good moods, consistent with prior research (Sanna et al., 1999) and with hypothesized gray arrows of Figure 1. Third, and of importance, contrasts indicated that when counterfactuals activated by moods and those suggested by self-motives were consistent, the net result was a quicker response. However, when counterfactuals activated by moods and self-motives were inconsistent, the net result was a slower response. This was shown for a wide range of self-motives.

STUDY 2: DISPOSITIONAL
SELF-MOTIVES AND OUTCOMES

Our second study was conducted with several goals. First, in Study 1, self-motives were manipulated directly. In Study 2, we took a more subtle approach to this issue. Instead of asking participants to consider motives, we preselected high-self-esteem (HSE) and low-self-esteem (LSE) persons, who differ in dispositional self-motives. HSE persons accept credit for success but deny blame for failure more than do LSE persons (Brown & Mankowski, 1993), and they can do this by using downward counterfactuals to mood-repair (Sanna et al., 1999). However, LSE persons use protective styles aimed at avoiding disapproval and defending identity (e.g., Arkin, 1981). Thus, HSE persons may be concerned with mood-repair and LSE persons may be concerned with self-protection.

Second, we manipulated outcome valence and repeatability. Sanna et al. (1999) found HSE persons to mood-repair in bad moods by using downward counterfactuals. However, participants did not believe they would have another try. Self-improvement may be salient on repeatable tasks. For example, although Markman et al. (1993) found downward counterfactuals predominated, upward counterfactuals were made after failures on repeatable tasks. Using repeatable tasks, we put mood-repair and self-improvement in opposition. We predicted HSE persons with no repeat after failure should agree with downward counterfactuals more and slower, replicating Sanna et al. (1999) but with generality to failures. However, for HSE persons with repeat after failure, self-improvement should be salient, and they should think of upward simulations more and quicker because self-improvement suggests a direction consistent with that activated initially after failure.

Other self-motives can be put in opposition. Sanna et al. (1999) found both HSE and LSE persons to think of downward counterfactuals more and quicker in good moods, perhaps due to mood-maintenance. For nonrepeatable tasks, this pattern should emerge after success. However, repeatable tasks again should make upward counterfactuals salient. Because upward is a direction inconsistent with that activated initially after success, counterfactuals should be agreed with slowly. Repeatability also is important to self-protection. LSE persons may “buffer” or “brace” for failure by using upward simulations (Sanna & Meier, 2000). Upward simulations used for self-protection may be particularly likely among LSE persons after failures when they expect to repeat. Study 2 used a 2 (self-esteem) × 2 (outcome) × 2 (repeatability) × 2 (counterfactual) mixed design.

Method

PARTICIPANTS

Participants were 103 (69 women and 34 men) students enrolled in an introductory psychology course who received extra credit. Approximately equal proportions of women and men were distributed among HSE and LSE groups and among conditions.

SELF-ESTEEM

Self-esteem was assessed using the Rosenberg Self-Esteem Scale (RSES) (Rosenberg, 1965). The RSES is a
well-validated measure of global self-worth. It consists of 10 items (e.g., “I take a positive view of myself”; “All in all, I am inclined to think I am a failure”) answered on 4-point scales (0 = strongly disagree to 3 = strongly agree). After appropriate reverse-scoring, a total is computed by summing (Brown & Mankowski, 1993). Fifty-three HSE ($M = 24.50$) and 50 LSE ($M = 15.25$) persons were classified from the upper and lower thirds of RSES scores from a larger pool of 276.

PROCEDURE

Procedures were similar to Study 1, except for preselection by self-esteem, manipulating outcome valence, and adding affect and preparation measures.

Outcome valence and counterfactuals. Outcomes were manipulated, instead of moods, to provide generality. Participants in the failure condition answered 10 items from the difficult RAT list; in the success condition, participants answered 10 items from the easy RAT list. These two lists are effective in inducing failure and success (see Mcfarlin & Blascovich, 1984). To augment list difficulty, outcomes were varied by providing (bogus) feedback (e.g., Sanna & Mark, 1995). After the 10 RAT items, participants read that they could calculate how well they performed. They read that because the RAT had been used in prior research, there were norms indicating how well they did in relation to others. They were told that these norms had been previously entered into the computer and they could calculate how well they did by pressing the spacebar.

Once the spacebar was pressed, there was a 9-s interval in which the screen flashed “CALCULATING… Please Wait.” Participants in the success condition read that they had performed very well and had scored in the top 20% of students tested at their university; participants in the failure condition read that they had performed very poorly and had scored in the bottom 20% at their university. Similar outcome valence manipulations using list difficulty plus bogus feedback have been used in prior research (e.g., Sanna & Mark, 1995).

Embedded among fillers, as an outcome manipulation check, participants answered two questions that asked them to rate the degree to which they thought they were successful on the RAT and the degree to which they thought their performance on the RAT was good (each anchored by 1 = not at all and 9 = very much). Counterfactuals were rated in a manner identical to Study 1.

Repeatability. Participants in repeat conditions were told that this would be the first in a series of two similar RATs that they would perform during the experiment. Participants in no-repeat conditions were told that after performing this one RAT, they would go on to perform an unrelated task that did not involve word associations.

Affect and preparation. Participants rated affect and preparation. The affect measure was identical to the mood-manipulation check in Study 1. The preparation measure consisted of three questions that asked participants to rate the extent to which they felt prepared, the extent to which they felt like they could handle, and the extent to which they felt ready to deal with another RAT task in the future (anchored by 1 = not at all and 9 = very much).

Results and Discussion

AGREEMENT AND LATENCY

$2 \times 2 \times 2 \times 2$ ANOVAs were conducted on agreement and latency, coded as in Study 1. For agreement, there was a marginal four-way interaction, $F(1, 95) = 3.82, p < .09$ (see Table 3), among Repeatability $\times$ Counterfactual, $F(1, 95) = 12.73, p < .01$; marginal Outcome $\times$ Counterfactual, $F(1, 95) = 3.79, p < .10$; and Self-Esteem $\times$ Outcome $\times$ Counterfactual, $F(1, 95) = 3.75, p < .10$, interactions. For latency, there were Repeatability $\times$ Counterfactual, $F(1, 95) = 9.03, p < .01$; Outcome $\times$ Counterfactual, $F(1, 95) = 5.81, p < .05$; and four-way, $F(1, 95) = 4.36, p < .05$, interactions. Hypotheses are most directly tested by contrasts within cells and between repeatability. Within no-repeat conditions, HSE persons after failure agreed with downward counterfactuals to a greater degree and more slowly than upward counterfactuals (agreement and latency $p s < .05$). Also within no-repeat conditions, HSE and LSE persons after success agreed with downward counterfactuals to a greater degree and more quickly (all agreement and latency $p s < .05$). These findings are consistent with past research (Sanna et al., 1999) and may indicate mood-repair and mood-maintenance motives.

Patterns differed in the repeat condition, however. This suggests qualifications to past research, consistent with our hypotheses. Within repeat conditions, HSE persons after failure agreed with upward counterfactuals to a greater degree and more quickly than with downward counterfactuals (agreement and latency $p s < .05$). Also within repeat conditions, HSE and LSE persons after success agreed with upward counterfactuals to a greater degree but more slowly (although LSE agreement differences were nonsignificant and latency was marginal, $p < .08$). In essence, the pattern of agreement is reversed compared with the corresponding cells of no-repeat conditions. Reaction times also were quicker for upward counterfactuals after failure and slower after success in these cells ($p s < .05$).

Together, findings are consistent with proposals. Expecting to perform a second RAT induced all participants to think about upward counterfactuals. Given that upward is a direction consistent with counterfactuals activated initially after failure, they were thought of
quickly among HSE persons with repeat conditions and produced a pattern different from no-repeat conditions. Upward is a direction inconsistent with that activated initially after success, and both HSE and LSE persons thought of them slowly, again producing different patterns between repeat and no-repeat conditions. LSE persons after failure thought of upward counterfactuals more and quickly (\(p < .05\)), irrespective of repeatability.

AFFECT AND PREPARATION

Affect was averaged as in Study 1 (Cronbach’s \(\alpha = .84\)), as were the three preparation questions (Cronbach’s \(\alpha = .77\)), with means presented in Table 3. Examining affect and preparation within each cell may help further distinguish among self-motives. For affect, a 2 (self-esteem) \(\times\) 2 (repeatability) \(\times\) 2 (outcome) ANOVA revealed main effects for self-esteem, \(F(1, 95) = 15.20, p < .01\), and a three-way interaction, \(F(1, 95) = 3.99, p < .05\). For preparation, there was a repeatability main effect, \(F(1, 95) = 18.83, p < .01\).

The repeatability main effect for preparation indicates that participants felt more prepared in all repeat cells than in all no-repeat cells (\(p < .05\)). Despite consistencies for preparation, affect diverged among conditions, suggesting that counterfactuals served different functions. With no-repeat, HSE persons felt good (but unprepared) after either failure or success, as did LSE participants after success (i.e., means for affect and preparation did not differ between these three cells). This is consistent with arguments that HSE participants used downward counterfactuals to mood-repair after failure, and both HSE and LSE persons preserved good moods (mood-maintained) after success.

With repeat, however, patterns in the corresponding cells reversed. That is, participants felt prepared but more negative, consistent with ideas that repeatability made self-improvement salient. Evidence for self-protection can be viewed in two ways: First, LSE persons after failure felt equally bad, irrespective of repeat, but felt more prepared with repeat (\(p < .05\)). Second, after failure with repeat, HSE and LSE persons felt equally prepared, but LSE persons felt worse (\(p < .05\)). In short, after failure, LSE persons’ affect is the same but preparation differs across repeatability, but within the repeat condition, LSE and HSE persons’ affect differs but preparation is the same. We speculate more about self-protection and possible reasons for this pattern in the General Discussion.

STUDY 3: REDUCED PROCESSING OPPORTUNITY AND OUTCOMES

A final study further explored people’s initial reactions. Sanna et al. (1999) found that upward and downward counterfactuals were a result of bad and good moods, respectively, when participants performed under time constraints. This provides some evidence for the hypothesized gray arrows in Figure 1. However, when mood-repair was an issue, people responded more slowly to downward counterfactual statements in bad moods, suggestive of a more effortful process. But initial responses to the range of self-motives outlined in Table 1 were not tested by Sanna et al. (1999), nor have they been examined by anyone else. Study 3 provided such a test.

Several studies find making judgments under time pressure results in less-effortful strategies (e.g., Bless, Mackie, & Schwarz, 1992; Mackie & Worth, 1989). For example, Mackie and Worth (1989) limited time available to participants considering persuasive messages and found reliance on heuristic cues (a less-effortful strategy) when forming judgments than without time limits. Smith and DeCoster (2000) also have noted that researchers have employed various “factors affecting cognitive capacity, whether in the form of time pressure, distraction from external stimuli or simultaneous tasks,
or resources such as task-relevant background knowledge" (p. 125).

To the extent that counterfactuals suggested by valenced events and self-motives are of an inconsistent direction, they may be reacted to slowly. A reduced opportunity to process may thwart effortful operations. We used a 4 (self-motive) × 2 (outcome) × 2 (time pressure) × 2 (counterfactual) mixed design, with counterfactual within-subjects. We predicted that low self-motives and more downward agreement and quicker latencies after failures, and more upward agreement and quicker latencies after successes, irrespective of self-motive. Study 3 also provides further generality to Study 1 by using outcome valence manipulations along with manipulated self-motives.

Method

PARTICIPANTS

Participants were 186 (108 women and 78 men) introductory psychology students who received extra course credit. There were approximately equal numbers, and approximately equal proportions of women and men, in each condition.

PROCEDURE

Procedures for Study 3 were similar to Study 1, with the following exceptions: First, outcome valence was manipulated instead of moods. Second, a reduced opportunity to process, in the form of time pressure, was manipulated in Study 3.

RAT and outcome valence. RAT and outcome manipulations were identical to Study 2.

Self-motives and counterfactuals. Four self-motives were manipulated, and counterfactuals were presented and rated, with instructions identical to Study 1.

Time pressure. In addition, unlike Study 1, half of the participants had a reduced opportunity to process counterfactual statements manipulated by time pressure. In the high-time-pressure condition, we imposed a time limit on participants’ responding (cf. Mackie & Worth, 1989). Participants read that they would have only 5 seconds to respond to each counterfactual statement. This time limit was determined on the basis of pilot testing to be minimally sufficient to read and respond to each statement; it also had been used in prior research (Sanna et al., 1999). As each statement was presented, a timer that counted off in 0.10-second increments also appeared on the top of the screen to let participants know how much time remained to respond. In the low-time-pressure condition, no time limits on responding were imposed, as had been the case in Study 1.

Results and Discussion

MANIPULATION CHECK

Two questions measuring participants’ perceptions of success on the RAT were averaged, r(184) = .72, p < .01, and were subjected to a 4 (self-motive) × 2 (mood) × 2 (time pressure) ANOVA. There was only a main effect of outcome valence, F(1, 170) = 15.33, p < .01 (M failure = 4.01, M success = 7.44), indicating that our manipulation was effective.

AGREEMENT AND LATENCY

4 × 2 × 2 × 2 ANOVAs were conducted on agreement and latency, coded as in our first two studies. For agreement, there was a four-way interaction, F(3, 170) = 9.93, p < .01 (see Table 4), and lower-order outcome, F(1, 170) = 5.91, p < .05; Self-Motive × Counterfactual, F(3, 170) = 6.32, p < .05; Outcome × Counterfactual, F(1, 170) = 16.39, p < .01; Self-Motive × Time Pressure × Counterfactual, F(3, 170) = 6.58, p < .05. For latency, there was time pressure, F(1, 170) = 15.25, p < .01; Outcome × Counterfactual, F(1, 170) = 6.68, p < .05; Time Pressure × Outcome × Counterfactual, F(3, 170) = 6.93, p < .05; and a four-way interaction, F(3, 170) = 7.31, p < .05.

Given our proposals, and the four-way interactions, tests of hypotheses are most clearly revealed by two 4 × 2 (self-motive) × 2 (outcome) × 2 (counterfactual) ANOVAs within low and high time pressure. The low-time-pressure condition is analogous to Study 1, and the pattern of means is similar to that study, seen by comparing the right half of Table 4 with Table 2. In the low-time-pressure condition, for agreement, there were interactions for Self-Motive × Counterfactual, F(3, 78) = 15.29, p < .01; Outcome × Counterfactual, F(1, 78) = 14.37, p < .01; and three-way interactions, F(1, 78) = 4.94, p < .05. Also under low time pressure, for latency, there was an Outcome × Counterfactual interaction, F(1, 78) = 14.22, p < .01. The pattern of means within these interactions was identical to Study 1, substituting failure and success for bad and good moods, respectively.

Most important, contrasts within each self-motive in the low-time-pressure condition also were similar to Study 1 and conform to predictions. For self-improvement after failure, upward counterfactuals were agreed with to a greater degree and more quickly than were downward ones (agreement and latency upward vs. downward ps < .05). Also for self-improvement, there was a similar pattern of agreement after success, but upward counterfactuals were agreed with more slowly (agreement and latency upward vs. downward ps < .05).

For mood-repair after failure, downward counterfactuals were agreed with more than were upward ones, and they were agreed with slowly (agreement and latency downward vs. upward ps < .05). Also for mood-repair, there was a similar pattern of agreement after suc-
cess, but upward counterfactuals were agreed with slowly
(agrernent and latency upward vs. downward ps < .05).
For mood-maintenance after success, downward
counterfactuals were agreed with to a greater degree
than were upward ones, and they were agreed with
quickly (for agreement and latency downward vs.
upward ps < .05). The reverse happened for mood-
maintenance after failure: Upward counterfactuals were
agreed with to a greater degree than were downward
ones, and they were agreed with quickly (agreement and
latency upward vs. downward ps < .05).

For self-protection after failure, upward counter-
factuals were agreed with more than were downward
ones and were made more quickly (agreement and
latency upward vs. downward ps < .05). Also for self-
protection was a similar agreement pattern after success,
but upward counterfactuals were agreed with slowly
(agreement, p < .10, and latency, p < .05, upward vs.
downward).

Replicating Study 1 under low time pressure, Study 3
is informative about initially activated counterfactuals
under high time pressure, because there was a com-
pletely different pattern. Under high time pressure, for
agreement, there was an outcome main effect, $F(1, 78) =
4.85$, $p < .05$, with a greater agreement after failure ($M =
6.47$) than success ($M = 5.65$). The only other effect was
an Outcome $\times$ Counterfactual interaction, $F(1, 78) =
15.17$, $p < .01$. Agreement was greater for upward ($M =
7.13$) than downward ($M = 5.81$) statements after failure
($p < .05$) and downward ($M = 6.46$) than upward ($M =
4.86$) statements after success ($p < .05$). As expected
given the imposed 5-s time limit, response latencies
within the high-time-pressure condition did not differ
from each other.

In short, Study 3 extends and elaborates our first two
studies and prior research, with several findings of note.
First, the pattern of means under low time pressure repli-
cated Study 1, here using outcome valence instead of
mood manipulations. Second, a completely different
pattern of mean agreement emerged in the high-time-
pressure condition. In all cases, upward statements elic-
ited greater agreement than did downward ones after
failure, but the reverse was true after success, a pattern
that supports the conjectured gray arrows in Figure 1
about initial responses.

GENERAL DISCUSSION

Our three studies provide converging evidence with
implications for understanding several processes under-
lying counterfactual thinking. As we discuss our results,
we once again find it heuristically useful to conceptual-
ize them in terms of relations depicted in Figure 1.

Outcomes and Moods as Antecedents
to Counterfactuals: Initial Reactions

Several studies demonstrated that outcome valence
(e.g., Markman et al., 1993; Roese & Olson, 1995; Sanna &
Turley-Ames, 2000) and mood (e.g., Sanna, 1998; Sanna
et al., 1998, 1999) can influence counterfactuals simi-
larly. After failures or bad moods, upward thoughts are
found; after successes or good moods, downward
thoughts are found. This is depicted by the gray arrows
in Figure 1 and was supported by the present results.
Moods and outcomes may produce similar effects due to
the informational value of feelings (Schwarz & Clore,
1996). Further support for this comes from findings that
moods do not influence counterfactuals when attributed
to irrelevant external sources (Sanna et al., 1998).
Upward counterfactuals were activated initially after fail-
uers (Studies 2 and 3) and bad moods (Study 1), whereas
downward counterfactuals were activated initially after

| TABLE 4: Mean Agreement and Response Latencies for Counterfactual Statements by Self-Motive, Outcome Valence, and Time Pressure for Study 3 |
|-----------------|-----------------|-----------------|-----------------|
|                 | Low Time Pressure |                | High Time Pressure |                |
|                 | Failure | Success | Failure | Success | Failure | Success | Failure | Success |
| Self-Motive     |         |         |         |         |         |         |         |         |
| Self-improvement |         |         |         |         |         |         |         |         |
| Agreement       | 8.21    | 4.86    | 7.62    | 6.00    | 7.92    | 5.98    | 5.13    | 6.73    |
| Latency(s)      | 5.20    | 7.69    | 8.00    | 4.82    | 4.83    | 4.80    | 4.97    | 4.95    |
| Mood-repair     |         |         |         |         |         |         |         |         |
| Agreement       | 5.01    | 7.43    | 4.35    | 6.19    | 7.21    | 6.20    | 5.31    | 6.59    |
| Latency(s)      | 6.32    | 8.43    | 7.34    | 5.38    | 4.27    | 4.34    | 4.80    | 4.82    |
| Mood-maintenance|         |         |         |         |         |         |         |         |
| Agreement       | 7.54    | 5.10    | 5.22    | 7.97    | 6.71    | 5.63    | 5.00    | 6.01    |
| Latency(s)      | 5.40    | 7.09    | 7.52    | 5.30    | 4.81    | 4.54    | 4.59    | 4.91    |
| Self-protection |         |         |         |         |         |         |         |         |
| Agreement       | 7.99    | 5.25    | 6.99    | 6.02    | 7.10    | 5.46    | 4.03    | 6.53    |
| Latency(s)      | 5.21    | 7.20    | 7.95    | 4.96    | 4.69    | 4.93    | 4.92    | 4.76    |
successes and good moods. Self-improvement may thus seem most salient after negatively valenced events and mood-maintenance after positively valenced events. The time-pressure results of Study 3 provide further evidence consistent with this pattern of initial activation.

We examined valenced events because they can contribute to an understanding of coping and well-being. Several studies, including our own (e.g., Sanna & Turley, 1996; Turley, Sanna, & Reiter, 1995), found more counterfactuals after negative events and quicker reactions (Roese & Hur, 1997). Some evidence for counterfactuals being more "automatic" after failures (Roese, 1997) was obtained in Study 3, under high time pressure. However, these prior studies did not assess counterfactual direction in response to both good and bad events. The present research demonstrates that downward and upward counterfactuals can occur more often and quicker in response to good and bad events, respectively (see also Sanna et al., 1999). The present research indicates that this pattern is moderated even further when various self-motives are made salient (Studies 1 and 3) or occur naturally (Study 2).

It is possible that particular patterns might not hold for every single study (cf. Roese & Olson, 1995), but we can speculate that exploring self-motives may hold a key to greater understanding.

**Self-Motives and Mental Simulations: Dual Processes?**

Research had focused on only self-improvement and mood-repair (e.g., Roese, 1994; Sanna et al., 1999; Taylor & Schneider, 1989), but other self-motives are theoretically and practically important, and by assessing them we provide a more complete picture of the functionality of mental simulations. Mental simulations can serve self-improvement, mood-repair, mood-maintenance, and self-protection. Sedikides and Strube (1997) classified mood-repair, mood-maintenance, and self-protection as alternate types of self-enhancement. People can repair, maintain, or protect a positive self-concept. These three types of self-enhancement may each serve an affective function. But of importance, they do so in different ways, and these ways had not been articulated. Mood-repair can be achieved by downward simulations after failure, mood-maintenance can be achieved by downward simulations after success, and self-protection can be achieved by upward simulations when anticipating performances. Perhaps an interesting question for future research is whether self-enhancement is a "master motive" that encompasses all others, a suggestion proposed but not tested (Sedikides & Strube, 1997).

More research on self-protection may be valuable. The affect and preparation of HSE and LSE persons in Study 2 suggests they may be using simulations in different ways (i.e., HSE felt good but LSE persons felt bad agreeing with upward counterfactuals in the repeat condition). HSE persons may be assimilating upward simulations for self-improvement, resulting in positive affect, but LSE persons are contrasting them for self-protection, resulting in negative affect (see Sanna, 1997, 2000). Repeat conditions of Study 2 also emphasize the past (counterfactual) versus future (prefactual) orientation of mental simulations. Although it may be functional for HSE persons to mood-repair after failure with no-repeat conditions (Sanna et al., 1999), it is not functional for HSE to mood-repair when they need to perform again. In fact, repeatability is inherent to self-improvement and self-protection. That is, one needs something to improve on, or protect from, in the future. Other motives such as self-verification and self-assessment (Sedikides & Strube, 1997) are also potentially relevant, but these similarly have received little attention in the counterfactual literature.

Finally, our research shares commonalities with dual-process models (e.g., Chaiken & Trope, 1999). In all three studies, responses were quicker to upward and downward counterfactuals when self-motives suggested a direction consistent with that activated initially. If directions suggested were inconsistent, responses were slower. Self-improvement may simply come to mind first, or most easily, after negative events or moods, whereas mood-maintenance may simply come to mind first, or most easily, after positive events or moods, as we described. Another account can predict the same initial reactions. Kahneman and Miller (1986) argued that counterfactuals recapitulate normality, and this involves heuristic processing. Upward and downward counterfactuals after negative and positive events, respectively, simply may be triggered as one contemplates a more neutral (normal) reality. Whether reactions result from motivational (self-motives) or cognitive (normality) processes may be worthy of further attention, as may be whether our findings are best characterized by selective, competitive, consolidative, or corrective (Gilbert, 1999) models. We hope that the present framework not only organizes and expands prior research but also affords new insights and provides an opportunity to test some novel and unique hypotheses.

**NOTES**

1. Participants answered an average of 5.68 Remote Associates Test (RAT) items in Study 1, which is about half of the control-list items. This is similar to prior research (McFarlin & Blascovich, 1984) and to our pilot testing, on which participants viewed performance quality in the absence of normative feedback or manipulated moods as relatively ambiguous (Sanna & Turley-Ames, 2000).

2. Ten upward (e.g., "I might have performed better on the RAT if only I had more time") and 10 downward (e.g., "I might have performed worse on the RAT if only I had less time") statements have been used in past research on a between-subjects basis (Sanna & Turley-
REFERENCES


