

5 Reflective and evaluative modes of mental simulation

Keith D. Markman and Matthew N. McMullen

A news item recently caught our attention. Flight attendant Kim Stroka claimed that she was too distraught to return to work after her co-worker died on United Airlines Flight 93, which was hijacked after taking off from Newark Liberty International Airport en route to San Francisco on 11 September 2001. Of compelling interest to counterfactual researchers, Stroka had apparently traded shifts with her co-worker and, thus, would have died instead of her colleague if she had worked her normal shift. Claiming that she was having difficulty eating and sleeping and that she was being treated by a psychologist for post-traumatic stress disorder, Stroka applied for medical and disability payments but was turned down by the state appellate court. According to the court, Stroka was not entitled to the award because “nothing happened while she was working which led to her current condition” (“No 9-11 Compensation for Flight Attendant,” Associated Press 2003).

A number of researchers have focused on the distinction between upward counterfactuals that simulate a better reality and downward counterfactuals that simulate a worse reality (e.g., Mandel 2003a; Markman *et al.* 1993; McMullen *et al.* 1995; Roese 1994; Roese and Olson 1995d; Sanna 1996, 2000). These researchers have adopted an approach that describes the possible functions that upward and downward counterfactual thoughts might serve. One function that has been identified is the contrast-based affective function (Roese 1997) – a given outcome will be judged more favorably to the extent that a less desirable alternative is salient. Thus, the strategic generation of downward counterfactuals may serve the function of enhancing coping and feelings of relative well-being by highlighting how the situation or outcome could easily have been worse.

Clearly, Kim Stroka has made a downward counterfactual. She did not die, but she can easily imagine how she could have died – indeed, she *would* have died. Just as clearly, however, generating this downward counterfactual has not made her feel any better. Instead, her consideration of the downward counterfactual world has engendered feelings of sadness, guilt, and fear. The Stroka case helps us make a more general point. Although contrast-based affective reactions to counterfactuals – whereby judgments are displaced

away from the counterfactual standard – are common, they are hardly the rule. Rather, Strack's downward counterfactual is *assimilative* in nature – her affective experience has been pulled *toward* the counterfactual standard (McMullen 1997).

In this chapter we will discuss the important aspects of a model (Markman and McMullen 2003) that attempts to explain how the very same counterfactual can engender dramatically different affective reactions. According to the model, the consequences of simulation direction are modulated by what we have termed *simulation mode* – relatively stronger tendencies to engage in reflective versus evaluative processing. In turn, we will describe how the interaction between simulation direction and mode produces important consequences for affect, motivation, and behavior.

The reflection and evaluation model

Reflection and evaluation

Markman and McMullen (2003) developed the Reflection and Evaluation Model (REM) of comparative thinking in order to provide an organizing framework for understanding how assimilation and contrast effects arise following counterfactual, social, and temporal comparisons. At the heart of the model is the assertion that two psychologically distinct modes of mental simulation operate during comparative thinking. The first of these modes is *reflection*, which is an experiential ("as if") mode of thinking characterized by vividly simulating that information about the comparison standard is true of, or part of, the self. The second of these modes is *evaluation*, which is characterized by the use of information about the standard as a reference point against which to evaluate reality (cf. Epstein *et al.* 1992; Oettingen 1996; Strack 1992).

Figure 5.1 depicts the interaction between simulation direction and mode. To illustrate, consider the student who receives a B on an exam but realizes that an A was easily attainable with some additional studying. In

Direction	Mode	
	Reflection	Evaluation
Upward	"I almost got an A"	"I got a B ... I failed to get an A"
Downward	"I nearly got hit by that truck"	"I was fortunate to not have been hit by that truck"

Figure 5.1 The interaction between simulation direction and mode.

the case of upward evaluation (UE), the student switches attention between the outcome (a grade of B) and the counterfactual standard (a grade of A). According to the REM, such attentional switching ("I got a B ... I could have gotten an A but instead I got a B") involves using the standard as a reference point and thereby encourages evaluative processing. In the case of upward reflection (UR), however, the student's attention is focused mainly on the counterfactual itself. According to the REM, focusing on the counterfactual encourages reflective processing whereby the student considers the implications of the counterfactual and temporarily experiences the counterfactual as if it were real ("What if I had actually gotten an A?"). In a sense, the student is "transported" into the counterfactual world (Green and Brock 2000; Kahneman 1995). Likewise, consider the case of a car driver who pulls away from the curb without carefully checking rear and side-view mirrors, and subsequently slams on the brakes as a large truck whizzes by. In the case of downward evaluation (DE), the driver switches attention between the counterfactual standard (being hit by the truck) and the outcome (not being hit by the truck), thereby encouraging evaluative processing ("I was fortunate to not have been hit by that truck"). In the case of downward reflection (DR), however, the driver's attention is mainly focused on the counterfactual itself, thereby encouraging reflective processing ("I nearly got hit by that truck").

Affect and the accessibility mechanism

Reflective processing and evaluative processing of counterfactuals yield predictable affective reactions and, according to the model, this is accomplished through an accessibility mechanism. Work by Mussweiler and his associates (e.g., Mussweiler 2003; Mussweiler and Strack 2000a, b) within the domain of social comparisons suggests that comparative self-evaluation produces two informational consequences. First, comparing oneself to a given standard increases the accessibility of standard-consistent knowledge about the self. Thus, upward comparisons render knowledge indicating a high standard of the self more accessible, whereas downward comparisons render knowledge indicating a low standard of the self more accessible. Second, and in turn, comparative self-evaluation provides a reference point against which the implications of this knowledge can be evaluated.

In a similar vein, we propose that counterfactual comparisons can also yield two informational consequences. First, making counterfactual comparisons should enhance the accessibility of cognitions about the self that are evaluatively consistent with the counterfactual standard. In turn, affect should be derived from thoughts about the standard that implicate the self, thereby yielding affective assimilation (Schwarz 1990; Schwarz and Clore 1983; Strack *et al.* 1985). To illustrate, consider an individual who learns that the aircraft she had originally planned to take crashed with everyone on board killed. Simulating the counterfactual possibility "I could have been on

that plane" (DR) renders standard-consistent cognitions about the self more accessible (e.g., "I could be dead," "I would never have been able to see my family again," "I would never have been able to accomplish what I wanted to in life"), and reflecting on these accessible cognitions produces counterfactual-congruent (in this case, negative) affect. On the other hand, employing the counterfactual as a standard against which to evaluate reality (DE) should produce positive affect via a contrast effect: "I'm lucky to be alive").

More generally, the notion that the very same counterfactual can produce both assimilative and contrastive reactions has intriguing implications for affective experience, as it may be that the mixed emotions (Larsen *et al.* 2001; Larsen *et al.* in press) that are often felt after events such as switching from the doomed plane flight are the result of reflective and evaluative modes of mental simulation operating in parallel (cf. Biernat and Manis 1994; Biernat *et al.* 1997; Mussweiler 2003; see also Markman and McMullen 2003 for a more detailed discussion of this issue). In this way, one can feel fortunate to be alive, yet deeply troubled by thoughts of what might have been.

Motivational consequences

In addition to the contrast-based affective function served by downward counterfactuals, counterfactual researchers have also focused on the preparative function that might be served by upward counterfactuals. Although upward counterfactuals may devalue the actual outcome and make us feel worse (e.g., Markman *et al.* 1993; Mellers *et al.* 1997; Roese 1994; Sanna 1996), simulating routes to imagined better realities may help us improve on our outcomes in the future (Johnson and Sherman 1990; Karniol and Ross 1996; Taylor and Schneider 1989). It has been suggested by some that counterfactual thoughts produce causal inferences (e.g., Hilton and Slugoski 1986; Lipe 1991; Wells and Gavanski 1989; but see also Mandel 2003c) and, according to Roese and his colleagues (e.g., Roese 1994, 1997; Roese and Olson 1997), it is this causal inference mechanism that underlies the preparative function. To illustrate, if Jim fails an exam, and then realizes that he would have passed if he had studied the textbook more carefully, he has identified an antecedent action that may trigger an expectancy regarding the consequences of taking that action in the future. In turn, this realization should heighten intentions to perform that action and thereby influence the production of that action.

The REM advances previous functional approaches, however, by suggesting that upward and downward counterfactuals can *both* have affective and preparative (or more generally, motivational) functions via reflective and evaluative processing. One of the key assumptions of the model is that the motivation to act, or not to act, is mediated by one's affective state, and also depends on the goal that has been adopted for performing a given task. Drawing on both Schwarz's (1990) feelings-as-information hypothesis (see

also Taylor 1991) and the mood-as-input perspective of Martin and his colleagues (e.g., Martin *et al.* 1993; see also Forgas 1995), the REM posits that negative affect should engender more persistence for tasks pursued to satisfy achievement goals (i.e., by employing the stop rule, "Have I done as well as I can do?") but lead to less persistence on tasks pursued merely for enjoyment (i.e., by employing the stop rule, "Am I still enjoying this task?"), whereas positive affect should engender more persistence for enjoyment tasks but lead to less persistence for achievement tasks. (See also Apter's (e.g., Apter 2001; Apter and Larsen 1993) distinction between telic and paratelic states.) Moreover, although the causal inference derived from the counterfactual comparison may suggest specific behaviors that one might perform in the future, we believe that the initial impetus to either change one's behavior or stay the present course is determined by affect. Overall, then, the REM specifies that affect and cognition make independent contributions to motivation and goal pursuit: affect motivates the individual to either change or maintain the status quo, whereas cognition shapes the specific strategies whereby one will either change or keep things the way they are.

Armed with this perspective on the influence of affect and cognition on motivation, specific predictions can be made regarding the motivational implications of upward and downward reflection and evaluation. To begin, UE yields negative affect, and should thus engender more persistence on achievement tasks but less persistence on enjoyment tasks. In addition, the causal inferences derived from UE (e.g., "I should have read the textbook chapters more carefully") should allow the individual to develop specific behavioral intentions and strategies regarding what actions should or should not be taken (see also Grieve *et al.* 1999; Morris and Moore 2000; Nasco and Marsh 1999). Expanding on previous functional approaches, however, the REM predicts that DR should also yield negative affect and thus exert effects on persistence similar to those produced by UE. Moreover, DR should produce causal inferences that seek to explain the event that almost happened (e.g., "I almost got hit by that truck because I didn't check my rear view mirror"). Thus, although DR does not help one to envision a *route* to a positive outcome, *per se*, we suggest that it can certainly motivate individuals to discontinue potentially destructive behaviors, in a manner not unlike fear communications that have been used in persuasion studies (e.g., Baron *et al.* 1994; Janis and Feshbach 1953).

On the other hand, DE engenders positive affect and may also yield causal inferences (e.g., "It's a good thing I studied as much as I did. If I had completely blown off my studying, I would have done much worse"). The causal inference derived here — that "some studying" is the cause of receiving a decent grade — indicates that a moderate amount of studying in the future will help maintain the status quo. Within the achievement domain, then, it is expected that the positive affect and specific causal inferences derived from DE will both contribute to an individual's complacency.

Expanding on previous functional approaches, however, the REM predicts that UR should also engender positive affect, as well as causal inferences that seek to explain the event that almost happened (e.g., "I almost sunk that thirty-foot putt because I accounted for the upward slope of the green"). Attempts to specify the nature of the influence of UR on motivation bring up two intriguing possibilities. On the one hand, it may be that the positive affect derived from UR will engender less persistence for achievement tasks (and more persistence on enjoyment tasks). On the other hand, the *realization* that one nearly accomplished the goal (e.g., making the putt), coupled with an understanding of *how* one nearly accomplished that goal (e.g., by accounting for the upward slope of the green), may instead engender feelings of self-efficacy (e.g., Bandura 1977; Sanna 1997) that empower one to persist and perform better on the task at hand. We will return to this issue a bit later.

In addition to producing emotions and suggesting causal inferences, counterfactual thinking may also influence regulatory strategies (Hur 2000; Pennington and Roese 2002; Roese *et al.* 1999, 2004). Higgins (1998) has argued that both promotion and prevention strategies are important means by which one can achieve desired end states. Promotion-oriented individuals, focused as they are on growth, advancement, and accomplishment, tend to pursue strategies aimed at approaching desirable outcomes, whereas prevention-oriented individuals, focused as they are on protection, safety, and responsibility, tend to pursue strategies aimed at avoiding undesirable outcomes.

Recent research has demonstrated how regulatory focus can be temporarily induced by cues in the environment (e.g., Forster *et al.* 1998, Higgins *et al.* 1997; Shah *et al.* 1998), and the salience of counterfactual standards may be one such situational cue. In this regard, an upward counterfactual represents a desirable outcome and thus may activate promotion goals in the service of obtaining that outcome, whereas a downward counterfactual represents an undesirable outcome and thus may activate prevention goals in the service of ensuring that the outcome does not occur (Hur 2000; Lockwood 2002; Lockwood *et al.* 2002).

According to the REM, the promotion focus activated by engaging in upward counterfactual thinking should play an important role in determining the nature of one's behavioral intentions. In the case of the student who failed to achieve an A, for example, a promotion focus should encourage the student to devise strategies designed to achieve favorable outcomes (e.g., putting more time into school work, attending class on a more regular basis). Conversely, the prevention focus activated by downward counterfactual thinking (e.g., "I almost got hit by that truck") should encourage the individual to adopt strategies designed to avoid bad outcomes (e.g., checking all rear view and side mirrors). Furthermore, and drawing once again on the feelings-as-information perspective, the REM predicts that prevention goals will be more highly activated after DR than after DE, and promotion

goals will be more highly activated after UE than after UR, because both DR and UE focus individuals on their *failure* to attain desired end-states.

Empirical tests of the reflection–evaluation model

Downward counterfactuals and motivation

Several laboratory studies have been conducted that test the REM's predictions in the domain of counterfactual thinking and motivation. To examine the motivational implications of DR and DE, McMullen and Markman (2000) measured students' responses after receiving their first exam grade in a course. All participants were instructed, in writing, to make a downward counterfactual (i.e., compare their present grade to an imagined worse grade). In the evaluation condition they were instructed to "evaluate your grade in comparison to the worse grade you imagined," whereas in the reflection condition they were instructed to "vividly imagine receiving that worse grade." Participants then indicated the extent to which they were experiencing various emotions and then answered several questions regarding their motivation to modify their study habits in the future (e.g., "How much do you feel you should change the way you study for the next exam?").

Consistent with predictions, more negative affect was experienced in the reflection condition, and more positive affect was experienced in the evaluation condition. Furthermore, motivation to modify future study habits was greatest in the reflection condition. Finally, mediational analyses indicated that the mode manipulation initially predicted motivation, but when affect was also entered into the regression equation, the mode coefficient dropped to nonsignificance, whereas the affect coefficient remained significant. Thus, and consistent with one of the key assumptions of the REM, affect mediated the counterfactual's impact on motivation.

Counterfactual thinking and task persistence

A second study (Markman *et al.* 2004b) examined the motivational implications of both downward and upward counterfactuals. Earlier in this chapter, we were equivocal with regard to predicting the effects of UR (e.g., feeling good by imagining having won the lottery) on motivation. If the feelings-as-information perspective is correct, UE should enhance motivation, whereas UR should lead to complacency. The prediction of a complacency effect following UR is supported by the work of Oettingen and her colleagues (e.g., Oettingen 1996; Oettingen and Mayer 2002; Oettingen *et al.* 2001). In these studies, engaging in positive fantasy by itself decreased motivation and inhibited success, whereas explicitly contrasting positive fantasies with reality enhanced motivation and facilitated success. According to Oettingen (1996), positive fantasies can be detrimental because they engender anticipatory consumption of motivation that would otherwise be directed toward achieving a given goal.

Markman *et al.* (2004b) gave participants as much time as they wished to solve an initial set (SET 1) of ten anagrams, with each anagram having multiple potential solutions. Following completion of SET 1, participants received "2X" feedback (e.g., if they found ten solutions, the computer informed them that they had found "ten out of the twenty possible solutions"), allowing them equal room to make either upward or downward counterfactuals. Participants in the UE and UR conditions were then asked to think about how their performance could have been better, with those in the DE condition being instructed to "compare their performance to the better performance they imagined," and those in the UR condition being instructed to "vividly imagine having performed better." On the other hand, participants in the DE and DR conditions were asked to think about how their performance could have been worse, and received reflection and evaluation instructions equivalent to the UE and UR participants. After generating their counterfactuals, participants responded to a set of mood adjectives, and were then given as much time as they wished to solve ten additional anagrams (SET 2). The dependent variables of interest were mood reports following SET 1 feedback, the amount of time they spent on SET 2 (persistence) relative to SET 1, and the number of anagram solutions correctly found in SET 2 relative to SET 1.

Analyses yielded the predicted direction by mode interaction for affect: URs reported more positive affect than did UEs, whereas DEs reported more positive affect than did DRs (see also McMullen 1997). Importantly, the predicted direction by mode interaction for task persistence was also obtained: whereas UEs persisted longer on SET 2 than did URs, DRs persisted longer on SET 2 than did DEs. In addition, UEs actually solved more SET 2 anagrams than did URs, although no differences were found between DRs and DEs. Finally, path analyses (see Figure 5.2) conducted on participants in the upward counterfactual conditions indicated that the relationship between mode (dummy coded: 1 = reflection, 2 = evaluation) and persistence was mediated by feelings of relaxation – the less relaxed participants felt, the more they persisted. In turn, SET 2 persistence predicted SET 2

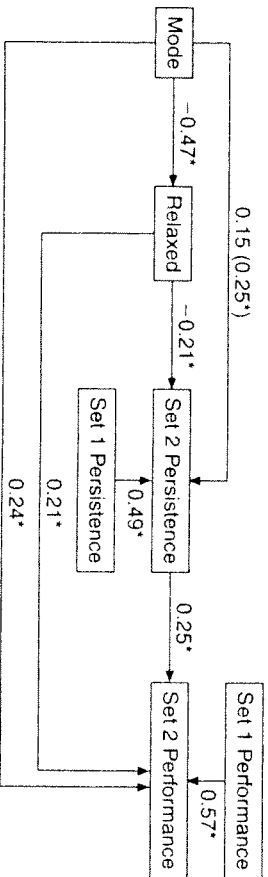


Figure 5.2 Mediator model of simulation mode, affect, and task persistence. Weights are standardized path coefficients adjusted for all other factors in the model. * $p < 0.05$.

performance after controlling for SET 1 persistence and SET 1 performance. Interestingly, however, an independent positive relationship also emerged between relaxation and SET 2 performance, suggesting that the affect derived from upward simulations can affect performance through two distinct mechanisms. On the one hand, UE may lead individuals to feel more aroused, thereby enhancing task persistence and task performance. Alternatively, however, the feelings of relaxation produced by UR may also enhance performance (despite the decrease in persistence), perhaps by facilitating the development of more creative solutions. In support of this possibility, a number of empirical studies have reported a relationship between positive mood and creativity (e.g., Hirt *et al.* 1997, Martin and Stoner 1996, Murray *et al.* 1990).

Although the analyses described above discovered a positive relationship between relaxation and performance, the independent negative relationship between UR and task persistence demonstrated by the Markman *et al.* (2004b) study supports Oettingen's (1996) notion that positive fantasies engender anticipatory consumption of motivation. According to Oettingen (1996: 238–9), in a positive fantasy,

a person may "experience" the future event ahead of time and may color the future experience more brightly and joyfully than reality would ever permit. Therefore the need to act is not felt, and the thorny path leading to implementing the fantasy may be easily overlooked.

Indeed, the counterfactuals generated by participants in the UR condition were characterized by this sort of flavor. For instance, one participant wrote, "I imagined the letters moving for me, instead of me going through them all individually and crossing them off in my mind. Meaning, I imagined the word appearing for me." This particular mental simulation may be optimistic, but it is also bereft of the implementation strategies (cf. Gollwitzer *et al.* 1990) that may be required to achieve the counterfactual outcome.

The complex relationships among UR, affect, motivation, and behavior can be further addressed by identifying and contrasting between two different types of counterfactuals. On the one hand, the UR participants in Markman *et al.* (2004b) engaged in what are essentially positive fantasies – they transported themselves into a better counterfactual world of *their own creation*. Kahneman and Varey (1990; Kahneman 1995; see also Teigen 1998b), however, have also described the special status of "close counterfactuals" (e.g., "John almost won the lottery," "Susan almost died"), which are characterized by a strong propensity for the counterfactual outcome to have existed soon before the actual outcome occurred. Propensities "... indicate advance toward the focal outcome, or regression away from it" (Kahneman and Varey 1990: 1105). Thus, to the extent that individuals perceive a trajectory toward either a desired or undesired state, assimilative effects

following counterfactual thinking are more likely to occur (see also Carver and Scheier 1990, Hsee *et al.* 1994, Landman and Percy 2000, Roese and Olson 1995b, Sanna *et al.* 2003, Tetlock 1998). Indeed, the casino game of Keno takes advantage of this phenomenon in a clever way: The numbers in the near vicinity of the winning number are lit up in addition to the number that won, giving rise to the feeling of "almost winning" (Sherman and McConnell 1995). Thus, close upward counterfactuals may sometimes encourage behavioral persistence.

Two of our studies have found evidence of affective assimilation following close counterfactuals. Markman and Tetlock (2000a) had participants engage in a simulated stock investment competition in which they chose between investing in one of two different companies. In the near-win condition, the chosen stock was just barely outperformed by the unchosen stock, whereas in the near-loss condition the chosen stock just barely outperformed the unchosen stock. After viewing the performance of the two stocks across a one-year span, participants indicated being happier when they nearly won (but lost) than when they nearly lost (but won). Similarly, McMullen and Markman (2002) found that fans of a basketball team that was losing by one point at half time but had come back from a substantial deficit felt better about the game than did fans of the team that was winning at half time (cf. Markman *et al.* 1995; but also Wohl and Enzle 2003 for an alternative perspective).

Reminiscent of Kahneman's (1995) distinction between elaborative and automatic mental simulations, we would suggest that while positive fantasies (i.e., elaborative simulations) of the quality described by Oettingen and her colleagues and elicited by Magkman *et al.* (2004b; see also McMullen 1997) may reduce motivation, close (i.e., automatic) upward counterfactuals that suggest that a better outcome was and, importantly, is plausibly attainable in the future, may actually serve to *increase* motivation. In fact, Markman and Tetlock (2000a) found initial support for this notion: Participants were *less* willing to reinvest in their chosen stock when they nearly lost (but won) than when they nearly won (but lost).

Counterfactuals, persistence, and goal type

Another experiment (McMullen *et al.* 2004) examined whether the influence of counterfactual thinking on motivation might also interact with the type of goal involved. Participants spent five minutes solving a set of crossword-like puzzles, and were then instructed to generate either a downward or upward counterfactual about their performance. Next, they were instructed to either vividly imagine the counterfactual (reflection) or to compare the counterfactual to their actual performance (evaluation). They then worked on another set of puzzles, but this time they could spend as much or as little time working as they wished. In the enjoyment condition, they were told that the point of the word puzzles was simply to have fun with the puzzles,

and if they were no longer having fun they could stop. In the achievement condition, on the other hand, they were told to perform as best they could, and when they were satisfied with their performance they could stop (cf. Sanna, Meier *et al.* 2001). The primary dependent variable was the amount of time they spent on the second set of word problems.

Consistent with the mood-as-input perspective (e.g., Martin *et al.* 1993), when participants engaged in the achievement task, UR reduced task persistence relative to UE, whereas DR increased persistence relative to DE. In the enjoyment task, however, this pattern was reversed. That is, UR actually increased task persistence relative to UE, whereas DR decreased persistence relative to DE. More generally then, motivation appears to result from rather complex interactions between simulation direction (upward versus downward), simulation mode (reflection versus evaluation), the affect produced by the simulation (positive versus negative), and the type of goal in question (achievement versus enjoyment).

Counterfactual thinking and regulatory focus

Finally, we have made an initial attempt to test the REM's predictions regarding the relationships between simulation direction, simulation mode, and regulatory focus. Within the domain of social comparisons, Lockwood *et al.* (2002) had participants compare themselves to positive versus negative role models and found that (upward) comparisons to the former were more apt to enhance academic motivation when promotion goals were primed, whereas (downward) comparisons to the latter were more apt to enhance academic motivation when prevention goals were primed. In the counterfactual thinking domain, Markman *et al.* (2004a) hypothesized that promotion goals would be most highly activated after UE because UE involves a concern with advancement (gains or the presence of positives), whereas prevention goals would be most highly activated after DR because DR involves a concern with safety (non-losses or the absence of negatives).

Participants recalled a somewhat negative event, made either an upward or downward counterfactual about it, and then engaged in either reflection ("imagine what might have happened instead") or evaluation ("think about the event and compare it to what might have happened instead"). Subsequently, participants completed Lockwood *et al.*'s (2002) strength of regulatory focus questionnaire, which includes items tapping the independent strength of both promotion and prevention foci (e.g., "I frequently think about how I can prevent failures in my life"), as well as a measure of academic motivation (e.g., "I plan to keep up with the reading assignments").

Intriguingly, the results indicated that UE and DR enhanced the strength of both promotion *and* prevention foci relative to the other conditions. In turn, UE and DR also enhanced intentions to study harder and improve upon one's academic performance. The difference between the present results and those of Lockwood *et al.* (2002) suggest that

counterfactuals may have more general motivational properties than do more specific and targeted social comparisons. We discuss potential differences between counterfactual and social comparisons at the end of this chapter.

Determinants of Reflection or Evaluation

Automatic versus controlled processing

Theorists have recently begun to examine the provocative question of what aspects of counterfactual thinking are more or less automatic versus controlled (cf. Bargh 1994; Kahneman 1995; Shiffrin and Schneider 1977). Roese (1997; Roese and Olson 1997; Roese *et al.* 2004) has argued that upward counterfactuals represent an automatic default in response to negative affect, whereas downward counterfactuals are constructed effortfully in an attempt to override negative affect. On the other hand, Sanna (2000; Sanna *et al.* 1999; Sanna, Chang *et al.* 2001) has suggested that upward and downward counterfactuals can be the result of either automatic or controlled processes, depending on the fit between outcome valence and/or mood and the most salient self-motive. Thus, negative outcomes or moods will automatically activate upward counterfactuals to the extent that self-improvement motives are salient, whereas positive outcomes or moods will automatically activate downward counterfactuals to the extent that mood-maintenance or mood-repair motives are salient. Conversely, a mismatch between outcome valence and/or mood (e.g., negative) and salient self-motive (e.g., mood repair) will instead stimulate the effortful construction of (in this case, downward) counterfactuals.

We are intrigued by these recent attempts to address this issue and would like to offer a new conceptual piece to the puzzle. The models by Roese and Sanna both assume that counterfactuals are initially and automatically *constructed* with reality (see also Gilbert *et al.* 1995; Wegener and Petty 1997). We would argue, however, that *assimilation* can also sometimes be the automatic default in counterfactual thinking.

To understand how assimilation can be the default, it is first important to consider what activates counterfactual generation in the first place. Roese (1997; Roese and Olson 1997) has argued that negative affect, construed as a response to goal blockage (Roese *et al.* 2004), is the “engine” that activates counterfactual thinking. We, however, prefer a broader conceptualization of how and why counterfactuals are generated. Drawing on Mandler’s (1964) notion that emotions occur in response to behavioral *interruptions*, we believe that counterfactual thinking may be automatically activated in response to one’s perception that an interruption has occurred in the natural order and flow of the behavioral “event stream.” Our notion of interruption is conceptually similar to Kahneman and Miller’s (1986) suggestion that counterfactuals are activated in response to violations of normality, but it broadens the normality notion by positing that counterfactuals will be automatically activated

in response to interruptions that are perceived in either the world that is (i.e., the actual event), or the world that could have been (i.e., the counterfactual event).

To illustrate, consider the student who typically receives A grades on exams but has this time received a B. For this student, a B represents an interruption in the typical event sequence (receiving A grades) and, thus, draws attentional focus. In this case, the grade that was will be automatically *contrasted* with the grade that could have been. This particular counterfactual would be categorized as an instance of UE, and we believe that UE follows from relatively automatic processing.

Perhaps even more interestingly, however, consider the case of the individual who switches from the doomed plane flight at the last minute, only to later learn that the plane has crashed, with all lives lost. Here, we would argue, attention is automatically drawn to the *counterfactual* because thoughts about what could have been (i.e., being killed on a plane flight) represent interruptions in the typical event sequence (i.e., surviving a plane flight). Thus, affective *assimilation* will be the default because one initially and automatically reflects on the counterfactual in the absence of any explicit comparison to reality. This counterfactual would be categorized as an instance of DR, and we believe that DR also follows from relatively automatic processing.

Conversely, we believe that DE and UR are driven by more effortful processes. As has been suggested by previous researchers (e.g., Markman *et al.* 1993; Roese 1994; Sanna 2000), DE is probably quite often an effortful attempt to maintain or ameliorate one’s present affect. Likewise, it probably requires some degree of effort to maintain an upward simulation while suppressing (cf. Wegner and Bargh 1998) potentially disturbing comparisons between the simulation and the real world. Indeed, although UR may be “cognitively easier” for the more fantasy-prone individual (Rhue and Lynn 1987), it is likely that some initial intent is required before even these types of individuals can become engrossed in their mental simulations of better possible worlds.

Motivational trade-offs

An important aspect of our work on counterfactual thinking has been our depiction of a critical tension between seemingly opposing motivations: future improvement versus affective enhancement (e.g., Markman *et al.* 1993; Markman and McMullen 2003; McMullen and Markman 2000). The road to future improvement via upward counterfactuals may engender inordinate amounts of negative affect (Roese and Olson 1997; Sherman and McConnell 1995), whereas the road to affective enhancement via downward counterfactuals may run the concomitant risk of engendering complacency and poor performance. In our view, the resolution of this preparative-affective trade-off plays an important role in determining whether reflection or evaluation will carry the day.

Indeed, this type of trade-off is not uncommon in psychology. For example, although overconfidence and unrealistic optimism may instantiate positive feelings, they may also be self-defeating (Weinstein and Klein 1995; Buehler *et al.* 1994; for a different view, see Taylor and Brown 1988). Likewise, positive fantasizing engenders positive affect at the expense of complacency and poor performance (Oettingen 1996). Research on self-handicapping suggests that people sometimes sabotage their prospects of success in order to enhance the availability of more comforting attributions about themselves and their abilities in the case of failure (Berglas and Jones 1978). In the realm of mental health, narcissists feel that they are superior individuals, yet often display self-defeating behavior patterns (Colvin *et al.* 1995), and perfectionists hold themselves to very high performance standards but put themselves at risk of depression (Blatt 1995). In a cross-cultural context, Americans exhibit higher academic self-esteem but they are outperformed by the more self-deprecating Japanese (Heine and Lehman 1999). Finally, cognitive dissonance theory asserts that one may either attempt to alter a negative behavior or rationalize it in order to maintain self-integrity (Aronson 1992).

This trade-off gives rise to a fundamental dilemma: Should I try to improve my performance, or should I try to feel better about myself? Although a variety of models, such as those just described, have addressed different aspects of this trade-off, the issue comes into particularly clear focus within the context of our reflection-evaluation model: UE and DR generally make people feel worse, but are motivating, while DE and UR generally make people feel good, but induce complacency.

What factors might determine which motive – future improvement or affective enhancement – an individual adopts? One factor that Markman *et al.* (1993; see also Sanna 1996, 1997) identified was whether or not an event was to be repeated in the future. When participants in this study believed that they were going to play future games, they were more likely to engage in upward evaluation – displaying a motivation to improve – but when they did not believe that they were going to play any more, they were more likely to engage in downward evaluation – displaying a motivation to feel good about what they have.

More generally, we suggest that perceptions of *attainability* are critical (Lockwood and Kunda 1997). According to the notion of the unidirectional drive upward (Festinger 1954), it is always preferable to successfully obtain a goal. However, when goal attainment is difficult or seemingly impossible, the only way to drive upward may be to fantasize – it is most often easier to *imagine* being a millionaire than to actually become one. We propose that when a goal is perceived as attainable, one is more likely to use comparative strategies that improve performance (i.e., UE and DR), whereas when a goal is perceived as unattainable, one is more likely to use comparative strategies that improve affect (i.e., UR and DE). A variety of other situational factors and individual differences undoubtedly play a role in determining the per-

ceived attainability of a goal, such as whether the event is to be repeated in the future (Boninger *et al.* 1994; Markman *et al.* 1993), the perceived probability of success (Teigen 1998b), feelings of self-efficacy (Sanna 1997), tendencies to engage in positive-constructive versus fear-of-failure daydreaming (Huba *et al.* 1981), incremental versus entity theories of intelligence (Dweck 2000), and differences in optimism versus pessimism (Scheier and Carver 1992). In turn, we would expect these factors to influence proclivities toward reflecting or evaluating.

One of the questions that this discussion raises is whether there must always be a trade-off between performance and affect. Are happy people destined to mediocrity, and successful people destined to depression? Is it not possible to feel both good *and* perform well? We suspect that both *are* possible. The strength of our reflection-evaluation approach is that the relationships among comparisons, affect, and motivation are not unidimensional; there are multiple avenues (i.e., simulation direction and mode) leading to both positive and negative affect, as well as to increased or decreased motivation. One fruitful avenue of investigation may be to search for asymmetries in comparisons. For example, it is possible that UR can be both motivating and affectively enhancing, whereas DR may *only* be motivating to the extent that it engenders negative affect.

Counterfactual versus social comparisons

We end our chapter with a brief discussion of differences between counterfactual and social comparisons. Researchers like Markman and McMullen (2003) and Olson *et al.* (2000) have focused on specifying common mechanisms underlying and common consequences accruing from counterfactual and social comparisons. However, we also believe that it is important for researchers to clarify what makes counterfactual comparisons *different* from other types of comparisons, and therefore worthy of the attention they have received.

First, we remind the reader of the findings obtained by Markman *et al.* (2004a), and how they were somewhat discrepant from those obtained by Lockwood *et al.* (2002). Whereas Lockwood *et al.* found that upward social comparisons enhanced academic motivation when promotion but not prevention goals were primed, while downward social comparisons enhanced academic motivation when prevention but not promotion goals were primed, Markman *et al.* found that UE and DR activated both promotion and prevention goals. We suggest that this occurred because counterfactuals may energize a broader class of motivations. Social comparisons typically focus on a *specific* target of comparison. In Lockwood *et al.*, this may have motivated participants to think about specific strategies whereby they could attain the outcomes experienced by the positive role model (i.e., promotion), or avoid the outcomes experienced by the negative role model (i.e., prevention). On the other hand, counterfactual comparisons are quite a bit more

diffuse – simulations are somewhat bounded by plausibility constraints but still remain fairly free to vary. Thus, participants in the Markman *et al.* study were free to imagine *many* or *multiple* ways whereby their academic outcomes could have been different. We speculate that the fewer constraints placed upon the choice or selection of counterfactual comparison targets encourages and allows the activation of more generalized and varied motivational strategies. Thus, after engaging in UE or DR following a negative academic event, an individual might pursue a desired end-state via both promotion (e.g., studying more) and prevention (e.g., reducing procrastination) means.

Second, we make note of potential differences between the mechanisms that underlie the processing of counterfactual and social comparisons. In a theoretical paper, Mussweiler (2003) attempts to specify how assimilation and contrast effects arise in comparisons. According to Mussweiler, the perceiver initially engages in a holistic assessment of the shared and unique features of the comparison referent and standard. With regard to social comparisons, if the perceiver decides that the self and the standard share common features, then the perceiver will test the hypothesis “How similar am I to the standard?” Testing for similarity will then heighten the accessibility of standard-consistent knowledge such that self-evaluations will be assimilated toward the standard. If, however, the perceiver decides that self and standard do not share common features, then the perceiver will instead engage in dissimilarity testing by asking, “How different am I from the standard?” In turn, testing for dissimilarity will heighten the accessibility of standard-inconsistent knowledge such that self-evaluations will be contrasted away from the standard.

Although the example described above focuses on social comparisons, Mussweiler (2003) suggests that similarity testing is the mechanism that accounts for assimilation and contrast effects in all types of comparisons. We agree that similarity testing may account for a wide range of social comparison phenomena. However, we also believe that it is substantially less useful for accounting for assimilation and contrast effects in counterfactual thinking. This can be illustrated easily enough by considering the student who just misses receiving an A in a class by a tenth of a percentage point. If the similarity-testing mechanism were applied, then the student would presumably arrive at the conclusion that her 89.4 was very similar to the 89.5 that she could have received and thus given her an A for the semester. According to Mussweiler, testing for similarity in this case should engender assimilation, thereby leading the student to feel good about her 89.4. Our intuitions, however, would suggest that the very opposite would occur – our student would probably be quite frustrated, bemoaning the fact that she “just missed” getting an A. Indeed, we would argue that it is the similarity of the real grade to the imagined grade that actually gives rise to such feelings of frustration (i.e., affective contrast)! After all, Kahneman and Tversky’s (1982b) participants judged that Mr Crane would be quite upset,

and not at all happy, when he discovered that his plane left just five minutes ago.

To illustrate a related point, consider Nike’s “Be Like Mike” (i.e., NBA former player Michael Jordan) advertising campaign a few years ago. Mussweiler (2003) suggests that individuals initially engage in a holistic assessment of one’s similarity to the standard and then engage in similarity or dissimilarity testing depending on the outcome of this initial assessment. Following this logic, most people should conclude that they are dissimilar to Jordan, thereby engendering contrast. However, our observation that children and adults alike pretend that they are “being like Mike” when they get on the basketball court indicates that this is clearly not what is happening. Instead, we argue that it is the act of reflecting on what it would be like to experience the success of Michael Jordan that enhances the accessibility of standard-consistent thoughts about the self – it is hardly necessary to test for similarity between the self and Michael Jordan in order to produce assimilation. In fact, testing the hypothesis that one is similar to Michael Jordan would simply highlight how dissimilar one is to him. Thus we believe that our reflection mechanism can account for these “Be Like Mike” effects in a way that Mussweiler’s hypothesis-testing mechanism cannot.

In sum, we hope that our reflection–evaluation approach stimulates novel hypotheses and opens a new window on counterfactual thinking that reveals more of its richness and complexity. In our view, no comprehensive approach to counterfactual thinking can succeed without incorporating assimilation and contrast effects. Many models of comparative thinking include assimilation and contrast as central components, whereas research on counterfactual thinking has lagged behind for substantially too long. By incorporating assimilation and contrast, we would argue, the affective and motivational issues that are so central to the functional approach to counterfactual thinking are fundamentally transformed: Counterfactual thoughts can motivate and discourage, assure and alarm, inspire and depress. In this chapter, we hope to have provided a glimpse into some of the new ideas that arise from this approach to counterfactual thinking.