Journal of Personality and Social Psychology

Accuracy and Artifact: Reexamining the Intensity Bias in Affective Forecasting

Linda J. Levine, Heather C. Lench, Robin L. Kaplan, and Martin A. Safer

Online First Publication, August 13, 2012. doi: 10.1037/a0029544

Journal of Personality and Social Psychology

0022-3514/12/$12.00 DOI: 10.1037/a0029544

Accuracy and Artifact: Reexamining the Intensity Bias in Affective Forecasting

Linda J. Levine
University of California, Irvine

Heather C. Lench
Texas A&M University

Robin L. Kaplan
University of California, Irvine

Martin A. Safer
The Catholic University of America

Research on affective forecasting shows that people have a robust tendency to overestimate the intensity of future emotion. We hypothesized that (a) people can accurately predict the intensity of their feelings about events and (b) a procedural artifact contributes to people's tendency to overestimate the intensity of their feelings in general. People may misinterpret the forecasting question as asking how they will feel *about* a focal event, but they are later asked to report their feelings in general without reference to that event. In the current investigation, participants predicted and reported both their feelings in general and their feelings about an election outcome (Study 1) and an exam grade (Study 3). We also assessed how participants interpreted forecasting questions (Studies 2 and 4) and conducted a meta-analysis of affective forecasting research (Study 5). The results showed that participants accurately predicted the intensity of their feelings about events. They overestimated only when asked to predict how they would feel in general and later report their feelings without reference to the focal event. Most participants, however, misinterpreted requests to predict their feelings in general as asking how they would feel when they were thinking about the focal event. Clarifying the meaning of the forecasting question significantly reduced overestimation. These findings reveal that people have more sophisticated self-knowledge than is commonly portrayed in the affective forecasting literature. Overestimation of future emotion is partly due to a procedure in which people predict one thing but are later asked to report another.

Keywords: affective forecasting, intensity bias, emotion, prediction

Supplemental materials: http://dx.doi.org/10.1037/a0029544.supp

People strive to attain outcomes that will make them happy and avoid outcomes that will make them miserable. The greater the intensity of emotion people expect an outcome to evoke, the more effort and resources they invest in attaining or avoiding it (Mellers & McGraw, 2001). Understandably then, accuracy in predicting future emotion has been called a virtual requirement for effective decision making (Loewenstein, 2007). Yet people's predictions about how they will feel have been found to be widely off the mark for outcomes as varied as the breakup of a relationship, being awarded or denied tenure, and the victory or loss of a favored political candidate or sports team (e.g., Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000).

How and why do our predictions go wrong? Research on affective forecasting shows that people tend to overestimate the impact that future events will have on their emotions—an error called the *impact bias* (Gilbert et al., 1998). One reason people make this error is *focalism* (Wilson et al., 2000) or the *focal illusion* (Schkade & Kahneman, 1998). At the time of prediction, people focus too much on salient features of a single future emotion-eliciting event. They fail to adjust their predictions sufficiently to account for the fact that there are sure to be other features and events that will also occupy their thoughts and influence their emotions. For instance, Wilson et al. (2000) found that college football fans overestimated how much their team's victory or loss would affect their happiness. This bias was mediated by fans' tendency to overestimate how much they would think about the outcome of the game; it was reduced, though not eliminated, by having fans consider a wide range of other events likely to occupy their thoughts after the game. People also fail to appreciate how quickly they will adjust to emotional events by regulating their emotions (Gilbert et al., 1998). Thus, the impact bias results in part from people focusing on salient information when predicting how they will feel and neglecting to consider co-occurring events and coping processes that will mitigate the intensity of their response (Gilbert & Wilson, 2007).

Linda J. Levine, Department of Psychology and Social Behavior, University of California, Irvine; Heather C. Lench, Department of Psychology, Texas A&M University; Robin L. Kaplan, Department of Psychology and Social Behavior, University of California, Irvine; Martin A. Safer, Department of Psychology, The Catholic University of America, Washington, DC.

We gratefully acknowledge the assistance of Danielle Jansen with data collection and preliminary analyses for Study 3.

Correspondence concerning this article should be addressed to Linda J. Levine, Department of Psychology and Social Behavior, University of California, Irvine, 4568 Social and Behavioral Sciences Gateway, Irvine, CA 92697-7085. E-mail: llevine@uci.edu

The impact bias is robust and persistent, and the contribution of focalism to this bias is well documented (e.g., Morewedge, Gilbert, & Wilson, 2005; Wilson et al., 2000). Nevertheless, people may have greater insight into their future emotional reactions than is conveyed by the affective forecasting literature. The impact bias actually encompasses several types of errors, though most forecasting studies do not distinguish between them (Wilson & Gilbert, 2003). People may overestimate how intensely they will feel about a future event—that is, the peak intensity of emotion they will experience immediately after an event occurs or when they are actively thinking about it later. People may also overestimate how much a future event will influence the intensity of their feelings in general—that is, their general emotional well-being irrespective of whether they are thinking about the event.

The goal of the current investigation was to clarify the nature of the impact bias, and mechanisms underlying it, by teasing apart people's ability to predict these two aspects of emotional intensity. Both of these aspects represent intensity bias, and it should be noted that this investigation does not directly address people's ability to predict the time course of their emotional responses, or duration bias.[1] We hypothesized that people can accurately predict the intensity of their feelings about events. We further hypothesized that, although people overestimate the impact of events on their general emotional well-being, the procedure commonly used to assess forecasting accuracy inflates the extent to which they overestimate. Thus, we were concerned with when, how much, and why people overestimate the intensity of these two concurrent aspects of emotional experience.

Can People Accurately Predict the Intensity of Their Feelings About Events?

According to prominent theories, emotions reflect people's experience of the relation between their goals or values and some set of circumstances in the world. Emotions are thus intentional states in the philosophical sense, meaning that they are *about* something (e.g., Clore & Huntsinger, 2009; Frijda, 1994; Solomon, 1973; Zeelenberg, Nelissen, Breugelmans, & Pieters, 2008). Although people sometimes feel cheerful, irritable, or anxious without being able to identify the cause or object of their feelings, more commonly, people feel happy about, angry at, or afraid of something. The layman's understanding of moods and emotions also reflects this distinction, with emotions viewed as having an identifiable source or object (Beedie, Terry, & Lane, 2005). Despite this, relatively few studies have asked people to predict and report their feelings specifically about a focal event. In those that have, the methods used have varied widely and yielded mixed results, sometimes showing overestimation (e.g., Buehler & McFarland, 2001; Eastwick et al., 2008; Mallett, Wilson, & Gilbert, 2008), and sometimes showing accuracy or underestimation (e.g., Böhm & Pfister, 2008; Crawford, McConnell, Lewis, & Sherman, 2002; Lench, Safer, & Levine, 2011). Thus, people's accuracy at predicting the intensity of their feelings about events remains an important unresolved issue.

People's feelings about events often differ dramatically from the effects of those events on their general emotional well-being. In one study, participants reported how they were feeling over several days using electronic diaries, both a few months before and a few months after September 11, 2001. A few months after the terrorist attacks, they also reported how they were currently feeling about the attacks. Ongoing reports of emotion did not differ before versus after the attacks, but participants reported feeling intensely anxious, sad, and angry about the attacks. The attacks thus elicited intense and lasting negative emotion, but these feelings were specific to instances when people were actively thinking about those events (Whalen, Henker, King, Jamner, & Levine, 2004). Similarly, the breakup of a relationship, being awarded or denied tenure, and the victory or loss of a favored political candidate may have little effect on people's day to day emotional experience (e.g., Gilbert et al., 1998), but people may experience intense emotion, even years later, when external circumstances or people's internal trains of thought bring those events to mind (Lench, Safer, & Levine, 2011; Levine, 1997; Levine, Safer, & Lench, 2006).

Accuracy in predicting both of these aspects of emotional experience is important. People make decisions in hopes of increasing their general emotional well-being. They also invest time, effort, and resources into pursuing and avoiding experiences that elicit emotion primarily while the experiences are occurring and when they are later thinking about them (e.g., musical performances, vacations, dentist visits, public speaking; Buehler & McFarland, 2001; Kahneman, Ritov, & Schkade, 2000). People may be far better at predicting the intensity of one aspect of emotional experience than the other, however. When predicting the impact of events on their feelings in general, people's tendency to focus on salient features of events should lead them to overestimate emotional intensity (Wilson et al., 2000). But this tendency may be appropriate, and lead to accuracy, when people predict how they will feel about events. Emotional arousal narrows the focus of attention to central features of events at the expense of peripheral features (for reviews, see Compton, 2003; Levine & Edelstein, 2009). So the features of emotion-eliciting events that are salient at the time of prediction are also likely to be salient when people are later thinking about those events. Therefore, we hypothesized that people would display fairly high accuracy when predicting the intensity of their feelings about events.

Does a Procedural Artifact Inflate Overestimation of Emotional Intensity?

People may also be better at predicting how they will feel in general than is commonly supposed. Focalism is a key mechanism

[1] This investigation does not address all of the errors encompassed by the impact bias. People may be poor at predicting the intensity of their emotional response, initially or after a delay. They may also be poor at predicting the time course of their emotional response, overestimating how long it will persist. A handful of studies have addressed the duration bias directly by asking people how often they will be in a good or bad mood after an event occurs (Wilson et al., 2000, p. 833; also see Igou, 2004) or by assessing predicted and experienced emotion at multiple time points and isolating bias in predicting emotional intensity from bias in predicting how quickly emotional intensity will diminish (Eastwick, Finkel, Krishnamurti, & Loewenstein, 2008; Hoerger, 2012; Kermer, Driver-Linn, Wilson, & Gilbert, 2006). Like the majority of affective forecasting studies, however, the current investigation addresses people's ability to predict the impact an event will have on the intensity of their emotional response at a particular point in time (e.g., days after the occurrence of a focal event), but does not directly address their ability to predict the time course of their emotional response.

underlying people's tendency to overestimate future emotion (e.g., Wilson et al., 2000). But another factor may also contribute to the magnitude and persistence of the intensity bias. Overestimation may be partly artifactual—an unanticipated consequence of the procedure most commonly used to assess forecasting accuracy. In everyday conversation, the context in which a question is raised influences how people interpret and respond to it (Grice, 1975). These conversational norms are also at work when participants interpret and respond to researchers' questions (Schwarz, 1999; Strack & Schwarz, 2007), and can influence the degree of bias demonstrated. For example, Ariely and Loewenstein (2000) presented evidence that duration neglect, people's tendency to ignore the duration of a past experience when rating how pleasurable it was, results in part from the operation of Gricean norms. They argued that participants interpret researchers' requests for a global evaluation of a past experience as a request to rate their average preference for the experience, which is typically considered independently of its duration. When participants were asked to evaluate experiences in contexts that reduced reliance on those conversational norms, they paid more attention to duration. With respect to affective forecasting, we hypothesized that the context in which forecasting questions are asked may lead participants to misinterpret these questions and promote overestimation of future emotion.

In a typical affective forecasting study, people are first asked to imagine that an event has occurred (e.g., the victory of a favored or disfavored political candidate), and to rate how they will feel in general after a specified period of time. After the event has occurred, and the period of time has elapsed, people are asked to rate how they are feeling in general without reference to the focal event. Thus, to ensure comparability, researchers ask people the same question about predicted and experienced emotion. Nevertheless, people may interpret these questions differently because of the contexts in which the questions are asked. People may overestimate the impact of events on their general emotional state because, in the context of having just been asked to imagine a specific future event, they interpret the forecasting question as asking how they will feel *about* that event, that is, when they are thinking about the event. When later reporting their emotional experience, however, the event is not mentioned and people may not be thinking about it. If people believe they are being asked to predict one thing, but are later asked to report another, inaccuracy is almost guaranteed.

To be clear, this is not focalism. People are displaying focalism when they understand that they are being asked to predict the impact an event will have on their general emotional state but they expect their general emotional state to be dominated by their reaction to the focal event. Because their general emotional state is instead influenced by a wide range of events and concerns, many of which are unrelated to the focal event, people tend to overestimate the intensity of their emotional response. We agree that focalism contributes to overestimating emotional intensity. However, we argue that the magnitude of overestimation is inflated by use of a procedure that leads people to misinterpret the general forecasting question. People may believe they are being asked to predict how they will feel when they are thinking about the focal event rather than how they will feel in general. Comparing these specific predictions with their later reports of how they are feeling in general would make it appear that people had overestimated the intensity of their emotional response when in fact they had misinterpreted the forecasting question. This view that forecasting studies might be comparing apples and oranges has been argued for and against (Lam, Buehler, McFarland, Ross, & Cheung, 2005; Wilson et al., 2000). It has never been tested, however, by contrasting procedures for assessing forecasting accuracy and examining how people interpret forecasting questions. We adopted this approach.

The Present Investigation

The present investigation addressed two fundamental issues in the affective forecasting literature. The first issue was whether the widespread claim that people overestimate emotional intensity is too broad: People may be accurate when predicting the intensity of their feelings about events but overestimate when predicting the intensity of their feelings in general. To test this, we directly compared the accuracy with which people predict how they will feel about a focal event versus how they will feel in general, using the same events and the same delay period. This comparison elucidates people's ability to predict two concurrent aspects of emotional intensity. The second issue was whether the procedure used most commonly to assess forecasting accuracy inflates the magnitude of the intensity bias, masking the sophistication of people's knowledge about their emotions. To address this issue, we assessed how people interpret general and specific forecasting questions. We also assessed forecasting accuracy using a procedure designed to clarify the meaning of the general forecasting question.

In Study 1, participants predicted and reported both their feelings in general and their feelings about the outcome of the 2008 U.S. presidential election. In Study 2, we assessed how participants interpreted these affective forecasting questions. In Study 3, participants predicted and reported both their feelings in general and their feelings about receiving an exam grade. We also assessed whether modifying the context in which the general forecasting question was presented increased forecasting accuracy. In Study 4, we assessed how participants interpreted these forecasting questions. In Study 5, we conducted a meta-analysis of affective forecasting research. We contrasted effect sizes for overestimation when studies assessed people's emotional experience in general without reference to the focal event versus when studies assessed people's feelings specifically about the focal event. This analysis also examined the timing of the question about emotional experience.

We tested the following hypotheses: (a) People should show high accuracy when asked to predict, and later to report, the intensity of their feelings about a focal event; (b) they should overestimate when asked to predict how they will feel in general, and later to report their feelings without reference to the focal event; (c) people misinterpret requests to predict their emotional state in general as asking how they will feel about the focal event, leading them to predict more intense emotion; and (d) clarifying the meaning of the general forecasting question should decrease the intensity of emotion predicted and increase forecasting accuracy.

Study 1: The Election

Study 1 assessed people's predicted and experienced reactions to the victory of Barack Obama over John McCain in the 2008

U.S. presidential election. We contrasted the effect on forecasting accuracy of asking people to predict and report their feelings in general versus asking people to predict and report their feelings specifically about the election outcome.

Method

Participants. Undergraduates in California, Texas, and Washington, DC ($N = 439$), completed online questionnaires before and after the election. The first questionnaire was completed 3 weeks before the election (October 8–14, 2008). On this questionnaire, most participants indicated a clear preference for Obama ($N = 238$) or McCain ($N = 156$) based on voting plans and candidate ratings. Analyses were conducted using data from these 394 participants (324 women, 70 men; $M_{age} = 19$ years, age range = 17–32 years), excluding those with no candidate preference.[2]

Design and procedure. Prior to questions concerning demographics and candidate preferences, we assessed predicted emotion. Participants were instructed, "Imagine that it is the week of November 4th, just days after the presidential election, and that Barack Obama won the election and will be the next President of the United States of America." Participants were randomly assigned to general versus specific predicted emotion conditions such that half were asked, "In general, how happy will you feel?" and half were asked, "How happy will you feel about Barack Obama being elected president?" Similar questions were asked about John McCain, and the order of questions about Obama and McCain was counterbalanced.[3]

The second questionnaire, completed 1–3 days after the election (November 5–7), assessed experienced emotion. Participants in each predicted emotion condition were randomly assigned to general versus specific experienced emotion conditions. Half were asked, "In general, how happy are you feeling these days?" without reference to the election outcome. Half were asked about their emotional reaction to the election outcome. Because we were interested in assessing feelings rather than attitudes, we specified, "By reaction, we mean the emotion that you are experiencing right now in response to the election. How happy do you feel about Barack Obama being elected President?" All emotional intensity ratings were made using a scale ranging from 1 (*not happy*) to 9 (*very happy*).

Results and Discussion

Participants' ratings of predicted and experienced happiness were analyzed using a mixed model analysis of variance (ANOVA). The within-subject factor was time (predicted emotion, experienced emotion). The between-subject factors were preferred candidate, predicted emotion question (general, specific), and experienced emotion question (general, specific). *F* values are shown in Table 1. Hedges's *g* was used as the effect-size measure for contrasts.

Not surprisingly, Obama supporters both predicted and experienced greater happiness with respect to Obama's victory than did McCain supporters. At the time of prediction, no main effect of question type was found (general vs. specific), and no interaction between candidate preference and question type was found. Thus, Obama supporters predicted that they would feel just as happy regardless of whether they were asked, "In general, how happy will you feel?" ($M = 7.43$, $SD = 1.68$) or "How happy will you feel about Obama being elected president?" ($M = 7.50$, $SD = 1.82$), $t(236) = .29$, $p = .99$, $g = .04$. McCain supporters predicted very little happiness. But again, their predictions did not vary depending on whether the forecasting question was general ($M = 3.01$, $SD = 2.14$) or specific ($M = 3.06$, $SD = 2.28$), $t(154) = .18$, $p = .99$, $g = .02$.

Were these predictions accurate or biased? As hypothesized, a significant interaction was found between candidate preference, the type of question asked about experienced emotion (general vs. specific), and time (predicted vs. experienced emotion). Figure 1 shows the mean intensities of predicted and experienced happiness for Obama and McCain supporters as a function of the question asked about experienced emotion. As Figure 1 (left side) shows, participants who were asked, "In general, how happy are you feeling?" without reference to the election, demonstrated the typical intensity bias. Both Obama and McCain supporters predicted more extreme emotional reactions than they later reported experiencing. Obama supporters, for whom the election outcome was positive, predicted that they would feel very happy but experienced more moderate happiness, $t(110) = 4.13$, $p < .001$, $g = .48$. McCain supporters, for whom the election outcome was negative, predicted very low levels of happiness but they too experienced more moderate happiness, $t(82) = -11.47$, $p < .001$, $g = 1.44$. So when the common procedure was used, both groups predicted a more extreme emotional reaction (positive or negative) than they later reported. In contrast, as Figure 1 (right side) shows, when participants were asked, "How happy are you feeling about Obama being elected president?" the intensity bias was reversed or eliminated, and participants were much more accurate. Obama supporters felt slightly, but significantly, happier than they predicted, $t(126) = -2.63$, $p = .01$, $g = .17$. McCain supporters both predicted and experienced little happiness, $t(72) = -0.73$, $p = .46$, $g = .05$.

[2] To allow us to assess forecasting accuracy for an outcome that was clearly positive or negative, analyses were conducted using data from participants who favored Obama or McCain. Most participants indicated on the preelection questionnaire that they planned to vote for Obama ($n = 195$) or McCain ($n = 124$). Participants also rated how good a president they thought Obama and McCain would be. Of those who were undecided, planned to vote for a third candidate, or were ineligible to vote, most indicated a preferred candidate by giving higher ratings to Obama ($n = 43$) or McCain ($n = 32$), so they too were included in analyses. We excluded data from participants who indicated no preference in voting plans or ratings ($n = 42$), or had missing data on emotion rating or candidate preference questions ($n = 3$). Results for all key analyses were the same as reported when analyses included all participants who completed both questionnaires (i.e., adding a group with no candidate preference).

[3] In Studies 1–4, we also assessed baseline general happiness ("In general, how happy are you feeling these days?") before referring to the focal event and before any experimental manipulation. This question, which is used in many affective forecasting studies, has been found to be highly correlated with more extensive scales of happiness and life satisfaction (for reviews, see Diener, 1984, p. 544; Gilbert et al., 1998, p. 620; Sandvik, Diener, & Seidlitz, 1993; Wilson et al., 2000, p. 823). As expected, baseline happiness ratings did not differ by experimental condition in any study.

Table 1
Analysis of Variance for Predicted and Experienced Happiness in Study 1

Source	*df*	*F*	*p*	η_p^2
Between-subject factors				
Preferred candidate	1	356.91	.001	.48
Predicted emotion question (T1 question)	1	0.99	.32	.00
Experienced emotion question (T2 question)	1	9.57	.002	.02
Candidate × T1 Question	1	0.64	.42	.00
Candidate × T2 Question	1	28.95	.001	.07
T1 Question × T2 Question	1	0.27	.61	.00
Candidate × T1 Question × T2 Question	1	0.11	.74	.00
MSE	386	(6.58)		
Within-subject factors				
Time	1	44.59	.001	.10
Time × Candidate	1	84.41	.001	.18
Time × T1 Question	1	1.60	.21	.00
Time × T2 Question	1	19.40	.001	.05
Time × Candidate × T1 Question	1	2.72	.10	.01
Time × Candidate × T2 Question	1	110.68	.001	.22
Time × T1 Question × T2 Question	1	0.35	.55	.00
Time × Candidate × T1 Question × T2 Question	1	1.49	.22	.00
MSE	386	(1.71)		

Note. Values enclosed in parentheses represent mean square errors (*MSE*s). T1 and T2 refer to Time 1 (predicted emotion) and Time 2 (experienced emotion), respectively.

We also assessed the correlation between predicted and experienced happiness when participants were asked a specific question versus a general question about their emotional experience. The results showed far greater correspondence when participants rated experienced emotion in response to a specific question, $r(198) = .89$, than a general question, $r(192) = .30$, $z = 10.96$, $p < .001$.[4]

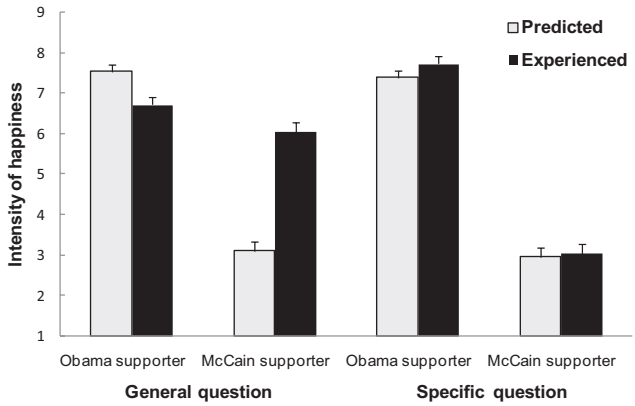

Figure 1. Study 1 means (and *SE*s) for predicted and experienced happiness for Obama and McCain supporters as a function of the type of question asked about experienced emotion. Use of the general question about experienced emotion led to the typical pattern of overestimating the emotional impact of a future event for both Obama and McCain supporters, whereas use of the specific question eliminated this bias. (In this figure, means were combined across the type of question asked about predicted emotion because predicted happiness did not differ depending on whether participants were asked a general or specific question.) Experienced happiness was assessed in two ways: general ("In general, how happy are you feeling these days?") and specific ("How happy do you feel about Barack Obama being elected President?").

In summary, the intensity of emotion participants predicted was just as extreme regardless of whether they were asked, "In general, how happy will you feel?" or "How happy will you feel about Barack Obama being elected president?" These predictions turned out to be overestimates only when participants were later asked how they were feeling in general without reference to the election outcome. Participants were much more accurate, and no evidence of overestimation was found, when they were asked specifically how they were feeling about the election outcome. These findings suggest that the claim that people overestimate emotional intensity may be too broad. As hypothesized, people overestimated the intensity of their feelings in general but showed high accuracy when predicting the intensity of their feelings about a specific event.

Study 2: Interpretation of Election Forecasting Questions

Why did participants' predictions in Study 1 turn out to be overestimates when they were later asked how happy they were feeling in general? Consistent with the idea of focalism in the affective forecasting literature, participants may have understood that they were being asked to predict their general emotional state. They may have overestimated anyway because they expected their general state to be determined primarily by their feelings about the election. Another possibility is that participants interpreted the request to predict how they would feel in general as asking how they would feel when thinking specifically about the election

[4] In Study 1, the correlation between predicted and experienced emotion of .30 is consistent with the results of a recent meta-analysis of affective forecasting research, which showed a mean correlation between predicted and experienced emotion of .28 across 16 studies (Mathieu & Gosling, 2012).

outcome—a misinterpretation that led them to predict intense emotion. To address this issue, in Study 2, we assessed how people interpret affective forecasting questions.

Method

Participants. Undergraduates at the University of California, Irvine, participated in the study ($N = 200$). They were recruited from a pool of students participating in social sciences research for partial course credit.[5]

Design and procedure. In an online questionnaire, participants were instructed to: "Imagine that it is the week of November 4th, 2012, three days after the next presidential election, and that the candidate you support won the election and will be the next President of the United States of America." Participants were randomly assigned to general or specific forecasting conditions. Those in the general forecasting condition ($n = 99$) were asked, "In general, how happy will you feel?" Those in the specific forecasting condition ($n = 101$) were asked, "How happy will you feel about the candidate you support being elected President?" Participants rated emotional intensity on a scale from 1 (*not happy*) to 9 (*very happy*).

The forecasting question that participants had just answered was then displayed at the top of the next page and participants were asked how they had interpreted it. They were instructed, "People interpret questions in different ways. Please help us understand how you interpreted this question." Participants selected one response from each of two pairs of alternative interpretations. We used alternative wordings, drawn from the affective forecasting literature, to convey general and specific interpretations so that the results would not hinge on a particular wording. From Pair 1, participants chose "I thought the question was asking approximately how happy I'll feel when the election outcome comes to mind" (specific) or "I thought the question was asking approximately how happy I'll feel most of the day" (general). From Pair 2, participants chose "I thought the question was asking approximately how happy I will be that the candidate I support won the election" (specific) or "I thought the question was asking approximately how happy my overall mood will be" (general). The order of the response options within each pair was counterbalanced.

Results and Discussion

As in Study 1, predicted happiness did not differ depending on whether the forecasting question was general ($M = 7.57$, $SE = 0.13$) or specific ($M = 7.58$, $SE = 0.12$), $t(198) = .10$, $p = .92$, $g = .01$. We examined participants' interpretations to find out whether they distinguished between the two forecasting questions. Not surprisingly, when presented with the first pair of alternative interpretations, the vast majority of participants who had been asked a specific forecasting question selected the specific interpretation (specific 89%, general 11%). Strikingly, the vast majority of participants who had been asked a general forecasting question also selected the specific interpretation (specific 75%, general 25%). The same pattern of results was found for the second pair of interpretations. Most participants who had been asked a specific forecasting question selected the specific interpretation (specific 91%, general 9%). But again, most who had been asked a general forecasting question also selected the specific interpretation (specific 81%, general 19%). For both forced-choice pairs, binomial tests indicated that significantly more participants selected the specific interpretation than the general interpretation, regardless of whether participants had been asked a specific question ($zs > 7.80$, $ps < .001$) or a general question ($zs > 4.92$, $ps < .001$).

Was misinterpreting the general forecasting question associated with predicting more intense emotion? To find out, we compared the intensity of emotion predicted by participants who had selected differing interpretations of the general forecasting question. For the first pair of alternative interpretations, participants who selected the specific interpretation predicted that they would feel happier ($M = 7.74$, $SE = .15$) than participants who selected the general interpretation ($M = 7.04$, $SE = .25$), $t(97) = 2.42$, $p = .02$, $g = .49$. The same pattern of results was found for the second pair of interpretations. Participants who selected the specific interpretation predicted that they would feel happier ($M = 7.69$, $SE = .14$) than participants who selected the general interpretation ($M = 7.05$, $SE = .29$), $t(97) = 1.96$, $p = .05$, $g = .40$.

These findings demonstrate that, as hypothesized, most participants misinterpret the request to predict their emotional state in general as asking how they will feel when they are thinking about the focal event. Moreover, those who misinterpret the question predict more intense emotion than those who correctly understand that they are being asked to predict their general emotional state. Thus, misinterpreting the forecasting question appears to contribute to people's tendency to overestimate the impact of events on their general emotional well-being.

Study 3: Not Making the Grade

Rather than reflecting a procedural flaw, misinterpreting the general forecasting question could itself stem from focalism (Wilson et al., 2000). That is, people's tendency to focus on peak emotional experiences, and neglect mundane experiences, may also lead them to interpret the general forecasting question as asking about their peak experience. If so, reducing the intensity bias would require directing people to consider mundane experiences when predicting how they will feel (Ayton, Pott, & Elwakili, 2007; Buehler & McFarland, 2001; Wilson et al., 2000). For example, Wilson et al. (2000) had students predict how they would feel in general in the days after their favored sports team won or lost an important competition. Overestimation was reduced by having students keep a detailed diary of the time they would spend on a wide range of activities during the days after the competition (e.g., going to class, socializing with friends, eating meals). Absent such painstaking measures, increasing forecasting accuracy is expected to be challenging (Wilson & Gilbert, 2005).

In contrast, we have argued that the context in which the general forecasting question is asked contributes to misinterpreting the question and inflates the intensity bias. If so, it should be possible to reduce this bias simply by changing the context in a manner that clarifies the meaning of the question. To test this, in Study 3, we asked undergraduates to predict and report their reaction to getting

[5] Due to administrative error, participant gender and age were not assessed for this study. Based on demographic information from three studies drawing from the same subject pool, and conducted within the same 12-month period, we estimate that about 85% of participants were female, with a mean age of 20 years and an age range of 17 to 39 years.

an exam grade that was lower than expected, expected, or higher than expected. We examined the extent to which they overestimated emotional intensity when asked the general forecasting question in the typical context (immediately after describing the focal event) versus in a "clarifying context." The clarifying context consisted of adding two questions designed to explicate the meaning of the general forecasting question.

In Study 3, we did not ask participants how they interpreted the general forecasting question in the typical versus the clarifying context. Doing so would likely have influenced participants' interpretations of, and responses to, subsequent questions. Therefore, we assessed how participants interpreted the forecasting questions in Study 4.

Method

Participants. University of California, Irvine, undergraduates, who were enrolled in an introductory psychology course, participated in the study for partial course credit ($N = 181$; 136 women, 40 men, five did not report gender; $M_{age} = 18.85$ years, age range = 17–28 years). We excluded nine participants who did not complete questions concerning predicted emotion, experienced emotion, or expectations concerning their grade, and we excluded four participants who had not checked their exam grade when they reported their emotional experience.

Design. Participants completed online questionnaires 2 weeks before their psychology midterm exam and 2 to 7 days after receiving their exam grade. Participants were randomly assigned to either the typical context condition or the clarifying context condition. Questionnaires were identical for these two conditions except that participants in the clarifying context condition completed two additional questions on the preexam questionnaire.

Procedure: Assessing predicted emotion. In the typical context condition, participants first reported the grade they expected to receive on their upcoming exam. Then they were asked a typical general forecasting question: "Suppose you get a grade that is lower than you expect. During the week after you find out your grade, in general, how happy will you feel?" Participants then rated how happy they would feel if they received the grade they expected, and if they received a higher grade than they expected. Finally, they answered demographic questions. All emotion ratings were made on a scale ranging from 1 (*not at all happy*) to 9 (*very happy*).

The questionnaire completed by participants in the clarifying context condition was identical except that two additional questions preceded the general forecasting question: (a) "During the week after you find out your grade, how happy will you feel about your grade?" (a specific forecasting question), and (b) "During the week after you find out your grade, do you think your grade will affect your overall mood?" (yes, no). Participants were then asked the same general forecasting question as in the typical context condition: "During the week after you find out your grade, in general, how happy will you feel?" Three considerations informed the inclusion of the additional questions. First, in this context, the general forecasting question was unlikely to mean "How happy will you feel *about* your grade?" Participants had just answered that question, so interpreting the general question as specific would have been redundant (Grice, 1975). Second, participants were not asked to exclude their reaction to their exam grade from consideration in predicting their general emotional state. Indeed, the second additional question explicitly invited them to consider whether their grade would influence their overall mood. Third, no attempt was made to counteract focalism by having participants consider a wide range of events in their lives that might mitigate the intensity of their emotional response (e.g., Wilson et al., 2000). The additional questions asked only about participants' reaction to their exam grade. Thus, the two additional questions were designed simply to clarify the meaning of the general forecasting question.

Procedure: Assessing experienced emotion. Two to 7 days after they received their exam grade, participants completed a second online questionnaire that assessed experienced emotion. This questionnaire was identical for the two experimental conditions. All participants were first asked, "In general, how happy are you feeling these days?" On a separate page, they then rated how they were feeling about the grade they had received on their exam. Emotion ratings were made on a scale ranging from 1 (*not at all happy*) to 9 (*very happy*). Finally, participants reported their exam grade.

Analyses. We first examined how happy participants predicted they would feel if they received a lower grade than expected, the grade they expected, and a higher grade than expected. We then examined the relation between predicted and experienced emotion for participants who actually received a lower, expected, or higher grade.

Results and Discussion

Predicted emotion. We examined predicted happiness in response to the general forecasting question using a mixed model ANOVA, with grade (lower than expected, expected, higher than expected) as the within-subject factor, and condition (typical context, clarifying context) as the between-subject factor. The results showed a main effect of grade, $F(2, 356) = 310.20$, $p < .001$, $\eta_p^2 = .64$, and an interaction between grade and condition, $F(2, 356) = 27.55$, $p < .001$, $\eta_p^2 = .13$. Participants in the typical context condition predicted that they would feel very little happiness if they received a lower grade than expected ($M = 4.33$, $SE = .21$), whereas participants in the clarifying context condition predicted a more moderate level of happiness ($M = 5.14$, $SE = .22$), $t(178) = 2.71$, $p = .01$, $g = .40$. Participants in the typical context condition predicted that they would feel very happy if they received the grade they expected ($M = 7.46$, $SE = .14$), whereas participants in the clarifying context condition predicted a more moderate level of happiness ($M = 6.78$, $SE = .15$), $t(178) = 3.38$, $p = .001$, $g = .50$. Participants in the typical context condition predicted that they would feel extremely happy if they received a higher grade than expected ($M = 8.44$, $SE = .13$), whereas participants in the clarifying context condition predicted a more moderate level of happiness ($M = 7.38$, $SE = .13$), $t(178) = 5.71$, $p < .001$, $g = .83$. Thus, participants who were asked, "In general, how happy will you feel?" in the typical context predicted that they would feel more extreme emotion (low happiness for a negative outcome, high happiness for positive outcomes) than participants who were asked the same question in a context that was designed to clarify its meaning.

We also compared the intensity of happiness predicted by participants who were asked the general forecasting question in the

typical context versus by participants who were asked the specific forecasting question ("How happy will you feel about your grade?"). We found no significant differences in predicted happiness for any grade outcome: receiving a lower grade (general: $M = 4.33$, $SE = .21$; specific: $M = 4.10$, $SE = .23$), an expected grade (general: $M = 7.46$, $SE = .14$; specific: $M = 7.13$, $SE = .15$), or a higher grade (general: $M = 8.44$, $SE = .13$; specific: $M = 8.13$, $SE = .13$; all ts < 1.70, all ps $> .09$). Thus, as in Studies 1 and 2, if the general forecasting question was asked in the typical context, predicted happiness did not differ significantly depending on whether the forecasting question was general or specific.

Predicted and experienced emotion in response to the general forecasting question. In true Lake Wobegon fashion, most students expected their exam performance to be above average (Dunning, Heath, & Suls, 2004). Only one student expected a grade lower than B−, and 60% of the students expected grades in the A range (A−, A, or A+). When grades were released, however, only 18% actually received grades in the A range. Thus, most students (77%) received a lower grade than they expected (72 in the typical context condition, 68 in the clarifying context condition); few received the grade they expected (11 in the typical context condition, 10 in the clarifying context condition) or a higher grade (11 in the typical context condition, nine in the clarifying context condition).

To examine predicted and experienced happiness in response to the general forecasting question for these three groups, we conducted a mixed model ANOVA. The within-subject factor was time (predicted emotion, experienced emotion) and the between-subject factors were received grade (lower than expected, expected, higher than expected) and condition (typical context, clarifying context). F values are shown in Table 2 and the mean intensities of predicted and experienced emotion are shown in Table 3. Because of the wide variation in cell sizes, post hoc comparisons between groups that received a lower, expected, or higher grade were conducted using the Tukey-Kramer adjustment.

As can be seen in Table 2, the results showed a significant main effect of grade, and significant interactions between grade and condition, between time and grade, and between time, grade, and condition. To explore these interactions, we first examined ANOVAs for predicted emotion and experienced emotion separately. Predicted emotion differed by grade, $F(2, 180) = 24.09$, $p < .001$, $\eta_p^2 = .22$, and a significant interaction was found between grade and condition, $F(2, 180) = 7.79$, $p = .001$, $\eta_p^2 = .08$. Comparisons of the means in Column A of Table 3 indicate that participants in the typical context condition predicted that they would feel less happy if they received a lower grade than if they received the grade they expected, $t(81) = -5.09$, $p < .001$, $g = 2.07$, or a higher grade, $t(81) = -6.37$, $p < .001$, $g = 2.49$. These participants also predicted that they would feel somewhat less happy if they received the grade they expected than if they received a higher grade, though this difference did not reach statistical significance, $t(20) = 2.67$, $p = .09$, $g = .46$. In contrast, comparisons of the means in Column C of Table 3 indicate that the more moderate emotional reactions predicted by participants in the clarifying context condition did not differ significantly by exam grade (all ts < 2.10, all ps $> .29$). Means for experienced emotion are shown in Table 3, Columns B and D. In contrast to predicted emotion, the ANOVA for experienced emotion showed no significant differences by exam grade or condition (all Fs < 1.29, all ps $> .26$).

Next we compared predicted and experienced emotion to find out whether participants showed an intensity bias. Means and paired comparison t tests are shown in Table 3. Participants who were asked the general forecasting question in the typical context showed a pronounced intensity bias. As shown in Columns A and B of Table 3, they predicted a more extreme emotional response than they experienced regardless of whether they received a lower grade than expected, the grade they expected, or a higher grade than expected. As shown in Columns C and D, the intensity bias was reduced or eliminated for participants in the clarifying context condition. Those who received a lower grade than expected predicted a more extreme emotional response than they experienced but the magnitude of the intensity bias was reduced. Indeed, the magnitude of the intensity bias (i.e., the absolute value of the difference between predicted and experienced emotion) was three times greater when the general forecasting question was asked in the typical context ($M = 1.94$, $SE = .26$) than when the same question was asked in the clarifying context ($M = 0.63$, $SE = .23$), $t(138) = 3.78$, $p < .001$, $g = .64$. Predicted and experienced emotion did not differ significantly for participants in the clarifying context condition who received the grade they expected or a higher grade. Thus, asking the general forecasting question in a

Table 2
Analysis of Variance for Predicted and Experienced Happiness in Study 3

Source	*df*	*F*	*p*	η_p^2
Between-subject factors				
Grade (lower, expected, higher)	2	9.25	.001	.10
Condition (typical context, clarifying context)	1	3.75	.05	.02
Grade × Condition	2	3.60	.03	.04
MSE	175	(4.61)		
Within-subject factors				
Time (predicted emotion, experienced emotion)	1	1.25	.27	.00
Time × Grade	2	22.00	.001	.20
Time × Condition	1	0.43	.51	.00
Time × Grade × Condition	2	6.36	.002	.07
MSE	175	(1.99)		

Note. Values enclosed in parentheses represent mean square errors (*MSE*s).

Table 3
Mean Predicted and Experienced Happiness for Participants Asked a General Forecasting Question in the Typical Context Versus in the Clarifying Context in Study 3

	Typical context condition			Clarifying context condition		
Exam grade	A. Predicted *M* (*SE*)	B. Experienced *M* (*SE*)	Comparison: paired *t*(*df*)	C. Predicted *M* (*SE*)	D. Experienced *M* (*SE*)	Comparison: paired *t*(*df*)
Lower than expected	4.42 (.22)	6.36 (.21)	$t(71) = -7.54^{***}$	5.40 (.23)	6.03 (.21)	$t(67) = -2.74^{**}$
Expected	7.81 (.57)	6.45 (.53)	$t(10) = 2.68^{*}$	6.00 (.60)	6.40 (.55)	$t(9) = -1.08$
Higher than expected	8.27 (.57)	6.55 (.53)	$t(10) = 3.30^{**}$	6.78 (.63)	5.44 (.58)	$t(8) = 1.71$

Note. Participants predicted how happy they would feel if they received an exam grade that was lower than expected, expected, and higher than expected. Those in the typical context condition were asked, "During the week after you find out your grade, in general, how happy will you feel?" Those in the clarifying context condition were asked the same general forecasting question except that it was preceded by two additional questions designed to elucidate its meaning. After receiving their grade, all participants rated experienced happiness in response to the question "In general, how happy are you feeling these days?"
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

context designed to clarify its meaning did not eliminate the intensity bias but significantly improved forecasting accuracy.

Predicted and experienced emotion in response to the specific forecasting question. Participants in the clarifying context condition also predicted and reported their feelings specifically about their grade. Here, no evidence of intensity bias was found (all ts < 0.84, all ps $> .40$). Participants who received a lower grade than expected predicted that they would feel very little happiness ($M = 4.34$, $SE = .26$), and later reported feeling just as little happiness as they had predicted ($M = 4.04$, $SE = .26$). Participants who received the grade they expected predicted that they would feel moderately happy ($M = 6.42$, $SE = .51$), and later reported feeling moderately happy ($M = 6.33$, $SE = .48$). Participants who received a higher grade than expected predicted that they would feel very happy ($M = 7.33$, $SE = .87$), and later reported feeling just as happy as they had predicted ($M = 7.78$, $SE = .55$). Thus, as a group, participants were highly accurate at predicting how they would feel specifically about their grade.

Correlations between predicted and experienced emotion. As another means of assessing forecasting accuracy, we computed correlations between predicted and experienced happiness. As hypothesized, the correlation between predicted and experienced emotion was higher when the general forecasting question was asked in the clarifying context, $r(85) = .48$, than in the typical context, $r(92) = .23$, $z = 1.92$, $p = .0549$. The correlation between predicted and experienced emotion in response to a specific forecasting question, $r(85) = .31$, did not differ significantly from either correlation in response to a general forecasting question.[6]

Discussion. The results of this study show that the magnitude of the intensity bias depends on the forecasting procedure used. When participants were asked how they would feel specifically about their grade, the correlation between predicted and experienced emotion was moderate. As a group, though, participants did not overestimate in response to the specific forecasting question. Indeed, intensity ratings were strikingly similar for predicted and experienced emotion for all grade outcomes: lower than expected, expected, and higher than expected. When a general forecasting question was asked in the typical context, participants showed a pronounced tendency to overestimate the extremity of their emotional response for all grade outcomes. When the same general question was asked in a context designed to clarify its meaning, participants who received the grade they expected or a higher grade did not overestimate emotional intensity. Participants who received a lower grade than expected did overestimate, but the magnitude of the intensity bias was reduced by two thirds. Correlations between predicted and experienced emotion also showed greater accuracy when the general forecasting question was asked in the clarifying context than in the typical context. Thus, asking the general forecasting question in a context designed to clarify its meaning improved forecasting accuracy.

[6] The correlation between predicted and experienced emotion in response to a specific forecasting question was not as high in Study 3 as in Study 1. Features of the focal event in Study 3 may account for this difference. Focal events (such as winning or losing an election or a game) rarely unfold precisely as anticipated (e.g., a person might anticipate a decisive victory but experience a win by a slim margin), but the amount of deviation depends in part on how the event is assessed. In Study 3, few students received the grade they expected or a higher grade, and getting a lower grade than expected encompassed a wide range of outcomes that differed markedly in their consequences and severity. For instance, among students who expected a B and received a lower grade, one may have received a B− and another may have failed the exam. To limit the length of the questionnaire, students were not asked to predict how they would feel about each of these distinct outcomes (e.g., "How will you feel if you receive a grade that is 1 level / 2 levels / 3 levels / etc. lower than you expect?"). The wide range of outcomes experienced within the category of "receiving a lower grade than expected" would tend to result in lower correlations between predicted and experienced emotion. Despite this feature of the design, the correlations were in the expected direction: somewhat higher in response to a specific forecasting question than in response to a general forecasting question in the typical context. We also examined correlations between predicted and experienced emotion using a revised classification system. Students who expected an A+ or A, and received a lower grade in the A range ($n = 6$), were initially classified as having received a lower grade than expected. Because they both expected and received a top grade, however, we also conducted analyses with these students classified as having received the grade they expected. Using this revised classification, a somewhat stronger correlation between predicted and experienced emotion was found for participants who were asked a specific forecasting question, $r(85) = .40$. The correlation between predicted and experienced happiness remained lower when the general forecasting question was asked in the typical context, $r(92) = .23$, than when the same question was asked in the clarifying context, $r(85) = .51$, $z = 2.13$, $p = .03$. For all other analyses, the results were the same as reported in the text.

Study 4: Interpretation of Grade Forecasting Questions

Study 3 showed that asking the general forecasting question in a different context decreased the intensity bias. In Study 4, we examined why this occurred. Did this change of context serve to clarify the meaning of the general forecasting question, resulting in more moderate predictions? Participants in Study 4 predicted how they would feel in general if they got a lower grade than they expected on an upcoming exam. As in Study 3, participants were asked the general forecasting question in the typical context or in a context designed to clarify its meaning. They were then asked how they had interpreted the question. Because this was likely to affect their interpretation of subsequent questions, participants in Study 4 were only asked how they would feel if they received a lower grade than expected, not a higher grade or the grade they expected. In addition, experienced emotion was not assessed.

Method

Participants. Undergraduates enrolled in an introductory psychology course at the University of California, Irvine, participated in the study for partial course credit ($N = 163$; 117 women, 46 men; $M_{age} = 19.08$ years, age range = 18–31 years).

Design and procedure. Students completed an online questionnaire 2 weeks before their midterm exam. As in Study 3, they were randomly assigned to a typical context condition or a clarifying context condition. Participants in the typical context condition ($n = 83$) were asked what grade they expected to receive on their upcoming exam. They were then asked a typical general forecasting question: "Suppose you get a grade on the midterm that is lower than you expect. During the week after you find out your grade, in general, how happy will you feel?" Participants in the clarifying context condition ($n = 80$) were asked the same questions, but the general forecasting question was preceded by the two additional questions described in Study 3: (a) "During the week after you find out your grade, how happy will you feel about your grade?" and (b) "During the week after you find out your grade, do you think your grade will affect your overall mood?" They were then asked the general forecasting question: "During the week after you find out your grade, in general, how happy will you feel?" Emotion ratings were made on a scale ranging from 1 (*not at all happy*) to 9 (*very happy*).

The general forecasting question was then displayed at the top of the next page and all participants were asked how they had interpreted it. They selected one response from each of two pairs of alternative interpretations. From Pair 1, participants chose "I thought the question was asking approximately how happy I will feel most of the time during that week" (general) or "I thought the question was asking approximately how happy I will feel when my midterm grade comes to mind during that week" (specific). From Pair 2, participants chose "I thought the question was asking approximately how happy my overall mood will be that week" (general) or "I thought the question was asking approximately how happy I will feel when I am thinking about my grade that week" (specific). The order of response options within each pair was the same for all participants: the general interpretation first, followed by the specific interpretation. Participants were also asked whether they took into account how they would feel about their midterm grade when predicting how they would feel in general. Finally, they answered demographic questions.

Results and Discussion

As in Study 3, participants who were asked the general forecasting question in the typical context predicted that they would feel worse ($M = 3.48$, $SE = .22$) than participants who were asked the same question in a context designed to clarify its meaning ($M = 4.94$, $SE = .17$), $t(161) = 5.22$, $p < .001$, $g = .81$.

In response to the first pair of interpretation options, when the general forecasting question was asked in the typical context, most participants misinterpreted it as asking how they would feel when thinking specifically about their grade (specific 63%, general 37%). When the general forecasting question was asked in a context designed to clarify its meaning, most participants correctly interpreted the question as asking about their emotional state in general (specific 39%, general 61%), $\chi^2(1, N = 163) = 9.31$, $p = .002$, $\varphi = .24$. A similar pattern of results was found for the second pair of interpretations. When the general forecasting question was asked in the typical context, over half of the participants selected the specific interpretation (specific 57%, general 43%). When the same question was asked in a context designed to clarify its meaning, most participants correctly interpreted the forecasting question as asking them how they would feel in general (specific 32%, general 68%), $\chi^2(1, N = 163) = 9.59$, $p = .002$, $\varphi = .24$.

In response to the two interpretation questions, 95 participants misinterpreted the general forecasting question by selecting the specific interpretation at least once; 68 participants selected the general interpretation in response to both questions, indicating that they clearly understood that they were being asked to predict how they would feel in general. To find out whether misinterpreting the general forecasting question was associated with predicting a more extreme emotional response (i.e., expecting to feel less happiness after receiving a poor grade), we conducted a regression analysis on forecast happiness. The predictors were question interpretation (specific for one or more questions = 0, general for both questions = 1), condition (typical context = 0, clarifying context = 1), and their interaction. The results showed that both misinterpreting the general forecasting question ($B = .86$, $SE = .42$, $\beta = .22$), $t(162) = 2.03$, $p = .04$, and being asked the forecasting question in the typical context ($B = 1.53$, $SE = .37$, $\beta = .40$), $t(162) = 4.13$, $p < .001$, were associated with forecasting a more extreme emotional response (less happiness). No significant interaction was found.[7]

We also conducted a mediation analysis using the method recommended by Preacher and Hayes (2008) to find out whether

[7] We assessed how participants interpreted affective forecasting questions in Studies 2 and 4 by having participants select a general or specific interpretation from two pairs of interpretation options. In both studies, by definition, the specific interpretation options referred to the focal event. As a result, they shared a few more words with the general forecasting prompt than did the general interpretation options. The additional shared words were *election outcome* (Study 2, Pair 1); *candidate, support, won, election* (Study 2, Pair 2); *midterm grade* (Study 4, Pair 1); and *feel, grade* (Study 4, Pair 2). These additional shared words may have encouraged some participants to select the specific interpretation. However, we also examined the intensity of emotion participants predicted before they could have been influenced by the wording of the interpretation options. As noted in the main text, in both studies, participants who misinterpreted the general forecasting question predicted more intense emotion than those who showed a clear understanding that they had been asked to predict their general emotional state.

question interpretation (specific for one or more questions vs. general for both questions) mediated the effect of condition on forecast happiness. This test was conducted using 5,000 bootstrapped samples to estimate the confidence intervals (CIs) associated with an indirect (mediated) effect of condition on forecast happiness. If the CI does not include zero, we can conclude that there is a significant mediation effect of condition on forecast happiness through question interpretation. The results showed that question interpretation partially mediated the relationship between condition and forecast happiness (mediated effect = .14, *SE* = .08, 95% CI [.02, .35]). In other words, the clarifying context condition reduced misinterpretation of the general forecasting question, which in turn reduced the extremity of forecast emotion. Finally, most participants reported that they had taken their grade into account in predicting their general emotional response regardless of whether they were in the typical context condition (88%) or the clarifying context condition (98%). Thus, the clarifying context did not appear to reduce predicted intensity by encouraging participants to exclude their grade from consideration.

These findings provide further support for the view that the procedure commonly used to assess forecasting accuracy promotes overestimation. They show that, when the typical forecasting procedure is used, most people misinterpret the general forecasting question as asking how they will feel when they are thinking about the focal event. When the same forecasting question is asked in a context designed to clarify its meaning, most people interpret it correctly as asking how they will feel in general. People who correctly interpret the general forecasting question predict a more moderate emotional response.

Study 5: Meta-Analysis

People overestimated emotional intensity when they were asked to predict how they would feel in general, and later to report their feelings without reference to the focal event—be it the election of President Obama or receiving an exam grade. Their predictions were much more accurate when they were asked to report how they were feeling *about* the focal event. To find out whether this pattern of results holds more broadly across different types of events and question wordings, we conducted a meta-analysis of affective forecasting research. We coded whether the question study participants were asked about experienced emotion was general or specific. We also coded the timing of the question about experienced emotion because it is likely to influence how people interpret the forecasting question. When people are asked how they are feeling immediately after the focal event occurs, it can be assumed that they are still thinking about the event, and that they will interpret the forecasting question as referring to their feelings about the event (Wilson et al., 2000). Therefore, we coded whether the question about experienced emotion followed immediately after the occurrence of the focal event or was delayed.

We expected the effect size representing intensity bias to be large when study participants were asked a general question about their emotional experience after a delay. We expected the effect size to be significantly smaller when study participants were asked a specific question about experienced emotion, either immediately or after a delay, or a general question immediately after the focal event.

Method

Study selection. Searches were conducted through the online database PsycINFO for articles published through the end of 2010 with the keyword *affective forecast**. Seventy-four potential articles were identified; additional articles were identified from their reference lists ($k = 28$). Articles were excluded if they (a) did not include empirical data, (b) did not report both forecasted and experienced emotion, (c) lacked information that would allow effect-size computation, (d) combined intensity ratings from general and specific questions so that it was not possible to extract separate effects for each type of question, (e) included an outcome other than emotional states (e.g., preference for a picture), or (f) were redundant with another source. Overall, 84 studies that met criteria were identified from 44 sources.[8]

Coding and effect extraction. To address our primary research question, we coded the question researchers asked about experienced emotion with respect to two factors: (a) question type: general or specific, and (b) question timing: immediate or delayed. Although these factors were coded separately, they were combined into a single moderator with four categories for analyses (general delayed, general immediate, specific delayed, specific immediate). This was done because, using meta-analytic techniques, interactions are estimated by contrasting levels of each group within a single moderator (Borenstein, Hedges, Higgins, & Rothstein,

[8] The articles listed below were reviewed and excluded for reasons described in the method section and labeled here. (a) The article did not include empirical data: Ariely, Huber, and Wertenbroch (2005); Elwyn and Miron-Shatz (2010); Johnson, Steffel, and Goldstein (2005); Loewenstein (1996, 2001, 2005, 2007); Loewenstein and Schkade (1999); MacInnis and Patrick (2006); Rachman and Bichard (1988); E. R. Smith and Mackie (2009); Zeelenberg, van Dijk, Manstead, and van der Pligt (2000). (b) The article did not report both forecasted and experienced emotion: Ariely and Loewenstein (2000); Bacova and Juskova (2009); Baron (1992); Bosson, Pinel, and Vandello (2010); Buehler, McFarland, Spyropoulos, and Lam (2007); DeWall and Baumeister (2006); Dillard, Fagerlin, Dal Cin, Zikmund-Fisher, and Ubel (2010); Falk, Dunn, and Norenzayan (2010); Gaunt, Sindic, and Leyens (2005); Geers and Lassiter (1999); Gilbert, Gill, and Wilson (2002); Golub, Gilbert, and Wilson (2009); Griffin, Dunning, and Ross (1990); Hartnett and Skowronski (2008); Igou (2004, 2008); Kassam, Gilbert, Boston, and Wilson (2008); Keller and Bless (2009); Loewenstein and Prelec (1993); Marshal and Brown (2006); Osberg and Shrauger (1986); Patrick, Chun, and MacInnis (2009); Samanez-Larkin et al. (2007); Schkade and Kahneman (1998); Seta, Haire, and Seta (2008); Sheldon, Gunz, Nichols, and Ferguson (2010); Ubel et al. (2001); Van Boven and Ashworth (2007); Van Boven, Loewenstein, and Dunning (2005); Walsh and Ayton (2009); Welsch and Kühling (2010); Wesp, Sandry, Prisco, and Sarte (2009); Wilson, Meyers, and Gilbert (2003); Winter, Moss, and Hoffman (2009); Wood and Bettman (2007). (c) The article lacked information that would allow effect-size computation: Dohke & Murata (2009, in Japanese); Dunn, Brackett, Ashton-James, Schneiderman, and Salovey (2007); Sheldon, Gunz, Nichols, and Ferguson (2010); Sweeny and Shepperd (2010); Totterdell, Parkinson, Briner, and Reynolds (1997); Wilson, Wheatley, Kurtz, Dunn, and Gilbert (2004). (d) The article combined intensity ratings from general and specific questions so that it was not possible to extract separate effects for each type of question: Eastwick et al. (2008). (e) The article included an outcome other than emotional states: Coughlan and Connolly (2001); Gilbert and Ebert (2002); Morewedge, Gilbert, Myrseth, Kassam, and Wilson (2010); Zhao and Meyer (2007).

2009). A question about experienced emotion was defined as *general* if it did not refer to the focal event (e.g., "How happy are you feeling?"). A question was defined as *specific* if it referred directly to the focal event (e.g., "How happy are you feeling about winning the game?"). A question was defined as *immediate* if the authors noted explicitly that the question followed immediately after the focal event (e.g., the participant won a game and was immediately asked, "How happy are you feeling?"). A question was defined as *delayed* if it did not follow immediately after the occurrence of the focal event. Any postponement (including a few minutes, intervening questions, or intervening procedures) was counted as a delay.[9] All of the ratings of predicted and experienced emotion that were analyzed were estimates of emotional intensity (e.g., "How happy will you feel?"), not estimates of emotion duration (e.g., "How long will your good mood last?"), although the intensity ratings differed with respect to the amount of time that had passed since the focal event occurred. Independent raters coded 40% of the studies in the meta-analysis for question type ($\kappa = .91$) and question timing ($\kappa = .93$). Discrepancies were resolved through discussion.

When experienced emotion was assessed multiple times, only the time point chronologically closest to the focal event was included. When authors' reports of experienced emotion were averaged over several time points, question timing was coded as delayed. In the supplemental materials available online, Table S1 shows the focal event for each study in the meta-analysis and the rationale that supported the coding of question type and timing.

The effect-size statistic Hedges's g was used in this study because it provides a more precise estimate of variance used in meta-analytic comparisons than Cohen's d (e.g., Hedges & Olkin, 1985; Rosenthal, 1991). Effect sizes of .2 are considered small, .5 medium, and .8 large. The program Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2005) was used to order, calculate, and compare effect sizes. When the information was available in text, table, or figure, effect sizes were calculated from means and standard deviations. If detailed information was not available, inferential procedures were used to estimate the effect using the comparison test statistic (typically a t test) and the associated p value. Lacking other information, if the effect was reported as nonsignificant, effect sizes were estimated by presuming $p = .50$. If the effect was reported as significant without additional information, effect sizes were estimated by presuming $p = .05$; this provides a conservative estimate because most actual p values would be smaller and associated with a larger effect size. These inferential procedures for effect-size estimation are standard practice in meta-analyses in order to include as many studies as possible in the review (Borenstein et al., 2009; Rosenthal, 1995). The results of all analyses remain identical, and effect sizes are within .02 units of those reported, if the six effect sizes derived using inferential procedures are removed.

CMA offers a correction to small-sample bias in computing effect sizes and this correction was applied (Borenstein et al., 2009). In addition, CMA was used to collapse across dependent effect sizes for relevant analyses when multiple outcomes or contrasts were assessed within the same study (Hunter & Schmidt, 2004). The precise formulae used to estimate effect sizes, weight effect sizes, and control for the dependency of effects are described in Borenstein et al. (2009; see also Lench, Flores, & Bench, 2011). It was typically not possible to determine the correlations among effect sizes from the original articles in order to calculate the combined variance for contrasts among groups; therefore, CMA assumptions were used because they tend to be conservative estimates. For ANOVA-type analyses, CMA presumes the correlation between outcomes in the same study is zero, which yields a conservative p-value estimate. For overall effect-size calculation, CMA presumes the correlation between outcomes is one, again yielding a conservative p-value estimate.

Random-effects models were calculated because variation in effect sizes among studies was assumed to occur as the result of random sampling error as well as differences between groups or individuals (Cooper & Hedges, 1994). The random-effects model therefore allows generalizations about the effects across a population rather than only to past studies (Raudenbush, 1994). The Q statistic was used to assess heterogeneity in the variance among the effect sizes. The categorical moderator was evaluated using the Q_{Between} (Q_{B}) statistic at the $p = .05$ level of significance. When significant, Q_{B} indicates that the effect sizes differ among the levels of the categorical moderator (Borenstein et al., 2009).

Analyses included 84 studies and 156 effect sizes. The total number of participants was 9,870. The results are shown in Table 4.

Results and Discussion

Effect sizes. The overall effect size was $g = .55$ (95% CI [0.42, 0.68]), $z = 8.29$, $p < .001$. This indicates that predicted emotion was typically greater than experienced emotion with a moderate effect size. A significant amount of variance was found, however ($Q = 606.73$, $p < .001$), suggesting that effect sizes were not homogeneous and the presence of moderators (Shadish & Haddock, 1994). Interactions in meta-analyses are estimated by contrasting levels within a moderator. Thus, in order to examine their combined effects, question type and question timing were included as one moderator of the difference between forecast and experienced emotion with four levels (general delayed, general immediate, specific delayed, specific immediate). Some studies manipulated factors hypothesized to reduce overestimation of emotional intensity (e.g., Andrade & Van Boven, 2010; Carlsmith, Wilson, & Gilbert, 2008). The approximate size of the effects, and the inferences drawn from analyses, do not change if these effect sizes (13) are omitted. Therefore, all studies were included in the analyses reported.

[9] In some studies, participants were asked filler questions after the focal event occurred but before rating their emotional experience. For example, participants in one study read a newspaper account of a child's death and were asked, "How well written was the story?" and "To what extent did the story keep your attention?" before reporting their feelings (Gilbert et al., 1998, p. 630). Instructing participants to attend to nonemotional features of events has been shown reliably to decrease the intensity of their emotional response to those events (e.g., Kalisch, Wiech, Herrmann, & Dolan, 2006). Because filler questions postpone assessment of experienced emotion, and may also serve to direct attention to nonemotional features of the focal event or to neutral events, we coded studies that included filler questions as delayed. In other articles, investigators did not explicitly specify the timing of the question about experienced emotion. The effect sizes from these studies were nearly identical to, and did not significantly differ from, the effects for delayed studies and were therefore included in the delayed category.

Table 4
Study 5 Meta-Analysis of Affective Forecasting Research

Study name	Study	Question type	Question timing	Hedges's *g*
Andrade & Van Boven (2010)	1	General	Immediate	−0.45
	2	General	Immediate	−0.32
Ayton, Pott, & Elwakili (2007)	1	General	Delayed	0.36
Böhm & Pfister (2008)	1	Specific	Delayed	−0.18
Buehler & McFarland (2001)	1	Specific	Delayed	0.61
	2	Specific	Immediate	0.32
	3	Specific	Delayed	0.80
Carlsmith, Wilson, & Gilbert (2008)	1	General	Delayed	1.14
	3	General	Delayed	0.83
Crawford, McConnell, Lewis, & Sherman (2002)	1	Specific	Immediate	−0.07
Dunn & Ashton-James (2008)	1	General	Delayed	1.37
	2	General	Delayed	0.60
	3	General	Delayed	1.16
Dunn, Biesanz, Human, & Finn (2007)	1	General	Immediate	0.03
	2A	General	Immediate	−0.25
	2B	General	Immediate	−0.31
	3	General	Immediate	0.70
	4	General	Immediate	0.50
Dunn, Wilson, & Gilbert (2003)	1	General	Delayed	1.01
	2	General	Delayed	1.00
Emanuel, Updegraff, Kalmbach, & Ciesla (2010)	1	General	Delayed	1.36
Fernandez-Duque & Landers (2008)	1	General	Immediate	−0.36
Finkenauer, Gallucci, van Dijk, & Pollmann (2007)	1	General	Delayed	0.87
Gilbert, Morewedge, Risen, & Wilson (2004)	1	General	Delayed	0.69
	2	General	Immediate	0.81
Gilbert, Pinel, Wilson, Blumberg, & Wheatley (1998)	1	General	Delayed	−0.10
	2	General	Delayed	0.69
	3	General	Delayed	0.67
	4	General	Delayed	0.46
	5	General	Delayed	0.95
	6	General	Delayed	1.06
Greitemeyer (2009)	1	General	Delayed	1.21
	2	General	Delayed	0.59
	3	General	Immediate	0.68
Hartnett & Skowronski (2010)	1	General	Delayed	1.25
Hoerger & Quirk (2010)	1	General	Delayed	0.79
Hoerger, Quirk, Lucas, & Carr (2009)	1	General	Delayed	0.84
Hoerger, Quirk, Lucas, & Carr (2010)	1	General	Delayed	0.34
Hsee & Zhang (2004)	1	General	Immediate	0.00
	3	Specific	Delayed	−0.36
Kawakami, Dunn, Karmali, & Dovidio (2009)	1	General	Delayed	2.15
	2	General	Delayed	0.78
Kermer, Driver-Linn, Wilson, & Gilbert (2006)	1	General	Delayed	0.77
	2	General	Delayed	0.54
Koo, Algoe, Wilson, & Gilbert (2008)	2, 3	General	Immediate	0.03
	4	Specific	Delayed	0.68
Ku (2008)	1	General	Delayed	0.93
	2	General	Delayed	0.62
	3	General	Delayed	0.91
Kurtz, Wilson, & Gilbert (2007)	1	General	Immediate	0.45
Lam, Buehler, McFarland, Ross, & Cheung (2005)	1	General	Delayed	0.66
	2	General	Delayed	0.56
Mallett, Wilson, & Gilbert (2008)	1	General	Immediate	0.62
	2	Specific	Delayed	0.58
Meyvis, Ratner, & Levav (2010)	1	General	Delayed	0.94
	2	Specific	Delayed	0.79
	4	General	Delayed	0.68
Morewedge, Gilbert, Keysar, Berkovitz, & Wilson (2007)	2, 4	General	Immediate	−0.27
	6	General	Delayed	0.52
Nielsen, Knutson, & Carstensen (2008)	1	General	Immediate	0.09
Pollmann & Finkenauer (2009)	1	General	Delayed	0.94
	2	General	Delayed	0.81
	3	General	Delayed	1.11

(*table continues*)

Table 4 (*continued*)

Study name	Study	Question type	Question timing	Hedges's *g*
Riis et al. (2005)	1	General	Delayed	0.15
Sanna & Schwarz (2004)	1	Specific	Immediate	0.45
Sevdalis & Harvey (2007)	1	General	Immediate	0.41
	2	General	Immediate	0.16
Sevdalis, Harvey, & Bell (2009)	1	Specific	Immediate	0.41
	2	Specific	Immediate	0.16
	3	Specific	Immediate	−0.58
Sieff, Dawes, & Loewenstein (1999)	1	General	Delayed	0.58
D. Smith et al. (2008)	1	General	Delayed	0.84
Tomlinson, Carmichael, Reis, & Aron (2010)	1	General	Delayed	0.85
	2	General	Delayed	1.67
	3	General	Delayed	1.43
van Dijk (2009)	1	General	Immediate	0.56
van Dijk, Finkenauer, & Pollmann (2008)	1	General	Immediate	0.31
Wilson, Centerbar, Kermer, & Gilbert (2005)	1	General	Delayed	1.23
	2	General	Delayed	−0.07
	3	General	Delayed	−0.46
Wilson, Meyers, & Gilbert (2003)	1	General	Delayed	0.88
Wilson, Wheatley, Meyers, Gilbert, & Axsom (2000)	1, 2	General	Delayed	0.71
	3	General	Delayed	1.10
Wirtz, Kruger, Scollon, & Diener (2003)	1	Specific	Immediate	0.00

Note. Effect sizes from the same study with the same moderator value were combined by hand for display in this table (the Comprehensive Meta-Analysis program combined effect sizes for all analyses). Two study numbers are indicated when predicted and experienced emotion for the same event were assessed in separate studies.

Analysis revealed a significant effect of the moderator, $Q_B(3) = 40.70$, $p < .001$. Follow-up analyses revealed that the effect size associated with general questions asked after a delay was large and differed from zero, indicating overestimation of emotional intensity (Hedges's $g = 0.80$, 95% CI [0.65, 0.95], $z = 10.52$, $p < .001$). This effect was significantly larger than the effect size associated with general questions asked immediately after the event (Hedges's $g = 0.16$, 95% CI [−0.03, .36]), $Q_B(1) = 26.27$, $p < .001$; specific questions asked after a delay (Hedges's $g = 0.37$, 95% CI [−0.04, .78]), $Q_B(1) = 3.70$, $p = .05$; and specific questions asked immediately after the event (Hedges's $g = -.01$, 95% CI [−0.28, .26]), $Q_B(1) = 26.14$, $p < .001$. The latter three effects did not differ significantly from one another. Effect sizes also did not differ significantly from zero when experienced emotion was assessed using general questions asked immediately after the event ($z = 1.64$, $p = .10$), specific questions asked after a delay ($z = 1.79$, $p = .07$), or specific questions asked immediately after the event ($z = -.05$, $p = .96$). Thus, studies that used these question types and timings did not, on average, demonstrate an intensity bias.

Assessing the likelihood of publication bias. Several methods were used to evaluate the likelihood that publication bias, the tendency for published studies available for meta-analyses to report significant results, influenced the reliability of the results. A funnel plot shows the treatment effect in relation to sample size and is visually examined for symmetry (asymmetry results when the size of the sample is related to the effect and indicates potential publication bias). A funnel plot for the present analyses appeared symmetrical, indicating that strong publication bias was unlikely. Classic fail-safe *n* calculation suggested 719 studies with nonsignificant effects would need to exist to threaten the reliability of the reported effect (Rosenthal, 1991). A more conservative calculation, Orwin's fail-safe *n,* indicated that the overall effect size in studies not included in the report would need to be .0000 to threaten the reliability of the effect (Orwin, 1983). Duval and Tweedie's (2000) trim-and-fill technique provided an observed overall effect size of .60 and an adjusted overall effect size of .60 (i.e., an estimate of the true effect size after accounting for the potential effects of publication bias), suggesting that publication bias was unlikely to influence the results.

Discussion. The results of the meta-analysis corroborate our experimental findings. Study participants showed a strong tendency to overestimate the intensity of future emotion when asked a general question about their emotional experience after a delay. Study participants were fairly accurate at predicting their feelings when asked a specific question about their emotional experience after a delay. They were also fairly accurate at predicting how they would feel immediately after the focal event, regardless of whether they were asked a specific or a general question about experienced emotion. Asking people how they are feeling specifically about the focal event, either immediately or after a delay, leads people to think about the focal event. People are also likely to be thinking about the focal event when asked a general question about their emotional experience immediately after the event occurs. Thus, the results of the meta-analysis are consistent with the view that study participants are fairly accurate at predicting the intensity of emotion they will feel when thinking about a focal event, even after a delay. Emotional intensity is overestimated when people are asked about their emotional experience without being reminded of the focal event either by the type of question asked or by its timing.

It is important to note that the effect size for overestimation was significantly greater for general than for specific forecasting questions, even when these questions were asked after a delay. This demonstrates that the type of question asked is important, not just its timing. Thus, the results of the meta-analysis cannot be explained solely in terms of a duration or decay bias in which people

are more accurate at predicting how they will feel immediately after an event than after a delay. However, these findings do not rule out the possibility that, in addition to question type, the length of the delay period also contributes to overestimation of emotion. Although it did not differ significantly from zero, the effect size when a specific question was asked after a delay was small to moderate. In addition, few studies asked participants a specific question about experienced emotion after a delay.

Limitations of this meta-analysis should also be noted. One limitation is that features of the focal events used in studies (such as their importance, controllability, and whether they occurred in the lab or in the field) may vary systematically with question type. No systematic variation of these factors with question type was evident in our review of the studies. Also, a strength of meta-analysis is the ability to combine across details of procedures. Nevertheless, this raises the possibility that study participants' greater accuracy when forecasting their feelings specifically about a focal event, as opposed to their feelings in general, might have been due to features of the focal event other than question type. In our empirical studies, however, we varied question type while holding the focal event (Obama's victory, an exam grade) constant. As in the meta-analysis, the results showed greater accuracy when people predicted their feelings about the focal event as opposed to their feelings in general. Taken together, these findings suggest that question type is not a proxy for unmeasured features of the focal events, but rather accounts meaningfully for differences in accuracy observed between groups in the meta-analysis.

A second limitation concerns variation in effect sizes within groups. The meta-analysis focused on differences in mean effect sizes between groups characterized by the type and timing of the question about experienced emotion. Within each group (e.g., specific immediate), studies varied with respect to the direction and magnitude of bias. A number of factors other than question type and timing may contribute to over- or underestimation of emotional intensity, including misconstruing the nature of the upcoming emotional event, empathy gap, and the importance of the focal event (for a detailed review, see Wilson & Gilbert, 2003). Though beyond the scope of the current investigation, these factors merit further study and may account for additional variation found in the direction and magnitude of bias within groups of studies.

General Discussion

Research on affective forecasting shows that people have a strong and persistent tendency to overestimate the intensity of their emotional reactions to future events (e.g., Ayton et al., 2007; Meyvis, Ratner, & Levav, 2010; Wilson & Gilbert, 2003; Wilson et al., 2000). In the present investigation, we extended research on the intensity bias in three ways: We directly contrasted people's ability to predict the intensity of two concurrent aspects of emotional experience—their feelings about an event versus their feelings in general; we examined how people interpret forecasting questions; and we conducted a meta-analysis of affective forecasting research. The results indicate that people can predict the intensity of their feelings about events with a high degree of accuracy, and that a procedural artifact contributes to people's tendency to overestimate the intensity of their feelings in general. Together, these findings reveal that people have a better understanding of their future emotions than is commonly portrayed in the affective forecasting literature.

People Can Accurately Predict the Intensity of Their Feelings About Events

In Study 1, people overestimated their reaction to the outcome of the 2008 U.S. presidential election when they were later asked to report how they were feeling in general without reference to the election. This bias was eliminated, and participants were much more accurate, when they were asked directly, "How happy are you feeling about Barack Obama being elected President?" Similarly, in Study 3, undergraduates overestimated their emotional reaction to their exam grade when later asked to report how they were feeling in general without reference to their grade. Overestimation was eliminated when they were asked directly how they were feeling about their grade. A meta-analysis of affective forecasting research showed that this pattern of results extends across a broad range of studies. The effect size representing overestimation was large when, after a focal event had occurred and time had passed, study participants were asked to report their emotional experience in general without reference to the focal event. The effect size representing overestimation was smaller, and did not differ significantly from zero, when study participants were asked to report their feelings specifically about a focal event, either immediately or after a delay. The effect size also did not differ significantly from zero when study participants were asked to report their emotional experience in general immediately after the focal event's occurrence, when they were likely to be thinking about the event.

This combination of the empirical studies and meta-analysis provides compelling evidence that people can predict the intensity of their feelings about events with a high degree of accuracy. Gilbert et al. (1998, p. 617) proposed that people may be pretty good at estimating the peak intensity of emotion they will feel immediately after a focal event occurs. We found, however, that accurate forecasting did not require immediate assessment of emotional experience. In Studies 1 and 3, people were asked to predict how they would feel, and later to report their feelings, days after learning the outcome of the election or their exam grade. Similarly, in the meta-analysis, the effect size representing overestimation did not differ significantly from zero when study participants reported how they were feeling specifically about a focal event after a delay. Partially consistent with Gilbert at al.'s view, then, our findings indicate that people are good at predicting the intensity of emotion they will feel when thinking about a focal event at the time it occurs, but they are also good at predicting their feelings about a focal event after a delay.

Several factors may contribute to this accuracy. Unlike when they predict their feelings in general, people have the same focal event in mind when they predict and report the intensity of their feelings about an event. Moreover, thinking about or experiencing an emotional event produces arousal that narrows attention to central features of the event (Compton, 2003; Levine & Edelstein, 2009; Safer, Christianson, Autry, & Osterlund, 1998). Thus, people's tendency to also focus on salient features of events at the time of prediction may promote accurate forecasts of their feelings about events. Finally, the procedures used in forecasting studies contribute to accuracy when people predict how they will feel

about events. People make distinctions between emotions and moods that are similar to those in the academic literature. In contrast to moods, they view emotions such as happiness as having an identifiable source or object (Beedie et al., 2005). In Study 2, about 90% of participants correctly interpreted the request to predict how happy they would feel about a focal event. In contrast, when asked to predict how happy they would feel in general, about 75% of participants misinterpreted the request as also asking about their specific emotional response. Thus, asking people how they will feel about an event appears to correspond to their intuitions about emotion, and correct interpretation of the forecasting question would favor accuracy.

The finding that people can accurately predict the intensity of their feelings about events has important implications (Zaki & Ochsner, 2011). It is widely accepted, as reflected in the research literature (e.g., Gilbert, Morwedge, Risen, & Wilson, 2004; Wilson et al., 2000), media reports (e.g., Gertner, 2003), and coverage of social psychology in modern textbooks (e.g., Gleitman, Reisberg, & Gross, 2010), that people can anticipate the valence of their emotional response to events but overestimate both emotional intensity and duration. The current findings demonstrate that this claim is too broad. They suggest instead that the victory or loss of one's favored candidate, entering and exiting relationships, attaining or not attaining tenure, often turn out to be just as euphoric or distressing as people expect them to be. These intense feelings can recur even years later when external events or internal trains of thought bring the events to mind (Carnelley, Wortman, Bolger, & Burke, 2006; Lench, Safer, & Levine, 2011; Whalen et al., 2004). Anticipated emotion informs people's decisions. The greater the intensity of emotion they expect an outcome to evoke, the more effort and resources they invest in attaining or avoiding it (Mellers & McGraw, 2001). Thus, people's accuracy in predicting the intensity of their feelings about specific events is an important finding that bodes well for effective decision making.

A Procedural Artifact Contributes to Overestimation of Emotional Intensity

In contrast to the accuracy with which they predicted their feelings about events, people showed a robust tendency to overestimate the intensity of their feelings in general. Focalism likely contributed to this bias but focalism alone cannot account for the findings. The results of the empirical studies show that the magnitude and consistency with which people overestimate emotion in affective forecasting studies are partly artifactual—a consequence of the procedure commonly used to assess forecasting accuracy.

We hypothesized that, in the context of having just been asked to imagine a specific future event, most people interpret the request to predict how they will feel in general as asking how they will feel about the event. Consistent with this view, in Studies 1–4, participants predicted that they would feel almost precisely the same intensity of emotion regardless of whether they were asked how they would feel in general or how they would feel specifically about the focal event. It appeared that they did not distinguish between these two questions. Studies 2 and 4 assessed people's interpretations of forecasting questions directly. In Study 2, 75% of participants misinterpreted the request to predict their emotional state in general as asking how they would feel when thinking about the election outcome. In Study 4, over 60% of participants misinterpreted the request to predict their emotional state in general as asking how they would feel when thinking about their exam grade. Importantly, participants who misinterpreted the general forecasting question predicted more intense emotion than those who correctly understood that they had been asked to predict their general emotional state.

Misunderstanding the forecasting question is not the same thing as focalism. People are displaying focalism when they understand that they are being asked to predict the impact an event will have on their general emotional state but expect that state to be dominated by their reaction to the focal event. Thus, focalism does not involve misunderstanding the forecasting question but rather giving the wrong answer to a question that was correctly understood. In contrast, our findings indicate that many people were never predicting their general emotional state at all.

If the context in which the general forecasting question is asked contributes to misinterpreting the question and inflates the intensity bias, it should be possible to reduce this bias simply by changing the context in a manner that clarifies the meaning of the question. In Study 3, students were asked to imagine receiving an exam grade and predict how they would feel in general. We compared the extent to which they overestimated emotional intensity when the general forecasting question was asked in the typical context (right after having been asked to imagine the focal event) versus in a context designed to clarify the question's meaning. In the clarifying context, participants were first asked how they would feel about their grade, and whether they thought their grade would influence their overall mood. They then predicted how they would feel in general. Based on the assumption that speakers strive to avoid redundancy (Grice, 1975), the additional questions were designed to convey that participants were being asked to predict how they would feel in general rather than how they would feel specifically about their grade.

The results showed that the magnitude of the intensity bias depended on the forecasting procedure used. When the general forecasting question was asked in the typical context, students showed a pronounced tendency to overestimate the intensity of their emotional response to receiving a higher grade than expected, their expected grade, and a lower grade than expected. When students were asked the same question in a context that clarified its meaning, they did not overestimate the intensity of their emotional response to receiving the grade they expected or a higher grade. They did overestimate their reaction to receiving a lower grade than expected, but the magnitude of the intensity bias was reduced by two thirds. Thus, clarifying the meaning of the forecasting question significantly improved forecasting accuracy.

These findings demonstrate that the concern that many forecasting studies are comparing apples and oranges is warranted. Participants' interpretations of research questions are strongly influenced by the content of adjacent instructions (e.g., Schwarz, 1999; Strack & Schwarz, 2007). Research can only inform us about the accuracy with which people can predict their general emotional state if research participants understand what it is they are being asked to predict. Our findings show that often they do not.

Alternative Explanations

Wilson et al. (2000) argued against the view that a procedural artifact contributes to overestimation in affective forecasting studies, so it is important to address their objections. They noted that "there is a sense in which this interpretation of our results is not an artifact but the point of the focalism hypothesis" (Wilson et al., 2000, p. 833). That is, when people make affective forecasts, they exaggerate how much they will think about the future emotional event. This may account for their tendency to interpret the general forecasting question as asking how they will feel when they are thinking about the focal event. Our findings show, however, that misinterpreting the general forecasting question is not an inevitable consequence of focalism. In Study 3, simply adding two questions sufficed to shift the majority of participants from misinterpreting the question to interpreting it as intended. It also led participants to forecast more moderate emotional responses. Thus, it is possible to design procedures less prone to misinterpretation to assess people's ability to predict the impact of events on their general emotional well-being.

We also conducted a study to find out whether misinterpreting the forecasting question can be accounted for by focalism or contributes independently to extreme affective forecasts (Levine & Kaplan, 2012). Modeling our study closely after Wilson et al.'s (2000) Study 4, we asked people to imagine that a tragic accident occurred at a space station a few months in the future, resulting in the deaths of several astronauts. Before reading about the tragedy, one group completed diary questionnaires that led them to think about events likely to occur during a typical day a few months in the future; the other group did not complete these questionnaires. All participants then predicted what their general level of happiness would be right after learning of the tragedy and over the next 3 days. The key modification introduced in our study was asking people how they had interpreted the request to predict their general level of happiness 2 days after the tragedy. We also asked how much they would think about the tragedy that day.

Consistent with Wilson et al.'s (2000) findings, people in the diary condition made less extreme affective forecasts for 2 days after the tragedy than did people in the no-diary condition. People in the diary condition were also more likely than those in the no-diary condition to interpret the general forecasting question correctly as asking about their overall mood rather than their feelings when thinking about the tragedy. Importantly, both expecting to think more about the tragedy and misinterpreting the general forecasting question contributed significantly and independently to more extreme forecasts.[10] These findings provide further evidence that, rather than being explained by focalism, misinterpreting the forecasting question contributes independently to more extreme forecasts in affective forecasting research.

Arguing against the view that question misinterpretation contributes to overestimation of the emotional impact of events, Wilson et al. (2000, p. 833) also cited studies in which people reported their emotional experience after being reminded about the focal event (Wilson et al., 1999) or moments after the event occurred (Gilbert et al., 1998, Study 6). People overestimated how unhappy they would feel even though these conditions made it likely that they were thinking about the focal event when they reported their emotional experience. The current investigation tested the extent to which people overestimate how they will feel when they are thinking about focal events using more stringent criteria. We examined studies in which, rather than providing a reminder during the procedure, the question about experienced emotion itself referred directly to the focal event (Study 1, Study 3, and meta-analysis). We also examined studies in which the question about experienced emotion followed immediately, rather than moments, after the focal event (meta-analysis). The contrast between the results of these procedures and the typical affective forecasting procedure was striking. Far greater forecasting accuracy was found when people predicted and reported the intensity of their feelings about, or immediately after, focal events.

Limitations

In summary, our findings show that people are fairly accurate at predicting the intensity of their feelings about events, and that a procedural artifact contributes to overestimation in predicting the intensity of their feelings in general. Limitations of this investigation should be noted, however. We do not mean to imply that there are no forecasting biases. This investigation addressed people's ability to predict the intensity, but not the time course, of their emotional responses. Even though people are fairly accurate at predicting how they will feel when an event comes to mind, even after a delay, they may still overestimate the pervasiveness of their emotional response and how long it will persist. For example, Wilson et al. (2000) found that people overestimated how much they would think about emotion-eliciting events and how often they would be in a good or bad mood after an event occurred. There is one way, however, in which our findings may apply to assessments of how quickly emotions wane over time. Studies of people's ability to predict emotional intensity

[10] Our methods and results are described here in more detail. As in Wilson et al. (2000), we asked participants ($N = 190$) to predict their affective response right after the focal event and over several days. We also adopted the procedure that Wilson et al. used in their Study 4 of having participants predict their general level of happiness using a scale that referred explicitly to average happiness as a reference point (1 = *below average happiness,* 5 = *average happiness,* and 9 = *above average happiness*). This allowed us to assess whether these procedures, used in a minority of affective forecasting studies, prevent misinterpretation of the general forecasting question. We focused on 2 days after the focal event (in the interpretation question, thinking question, and analyses) because Wilson et al. found that the effects of the diary manipulation were strongest 2 or 3 days after the focal event. The results showed that people in the diary condition made less extreme affective forecasts for 2 days after the tragedy ($M = 4.34$, $SE = .13$) than did people in the no-diary condition ($M = 3.96$, $SE = .13$), $t(288) = 2.06$, $p = .04$, $g = .24$. People in the diary condition were also more likely (68%) than those in the no-diary condition (56%) to interpret the general forecasting question correctly as asking about their overall mood that day (general) rather than their feelings when thinking about the tragedy (specific), $\chi^2(1, N = 290) = 5.15$, $p = .02$. A regression analysis examining forecast happiness 2 days after the accident showed that both expecting to think more about the accident ($B = -.28$, $SE = .04$, $\beta = -.36$), $t(289) = -6.80$, $p < .001$, and misinterpreting the general forecasting question as specific ($B = .78$, $SE = .17$, $\beta = -.24$), $t(289) = -4.58$, $p < .001$, contributed significantly, and independently (β for the interaction $= -.06$, *ns*), to more extreme forecasts. In summary, in a minority of affective forecasting studies, people have been asked to predict their feelings over several days or to predict their feelings using average happiness as an explicit reference point. These procedures may have reduced somewhat, but did not prevent, misinterpretation of the general forecasting question. Focalism and misinterpreting the forecasting question both contributed to more extreme affective forecasts.

and emotion duration often use the same basic procedure (e.g., Finkenauer, Gallucci, van Dijk, & Pollmann, 2007; Gilbert et al., 1998). Misinterpreting the forecasting question may also contribute to overestimation of emotion duration if people predict how they will feel when an emotional event comes to mind after each of a series of days, weeks, or months, but are later asked to report their emotions in general at each time point.

We also do not mean to imply that people are always accurate when predicting emotional intensity. In Study 3, for example, the magnitude of the intensity bias decreased significantly when the meaning of the general forecasting question was clarified, but students still overestimated the impact that a poor exam grade would have on their general emotional well-being. Overestimation has also been demonstrated using alternative procedures. For example, people overestimate when predicting the average intensity of emotion they will experience over a period of time, such as during a vacation (Wirtz, Kruger, Scollon, & Diener, 2003). Researchers have shown convincingly that focusing on moments of peak intensity at the time of prediction contributes to overestimation, and that factors that moderate focalism can decrease overestimation (e.g., Morewedge et al., 2005; Wilson et al., 2000). But people can also be taken aback by the unexpected intensity and pervasiveness of their emotional responses (e.g., Lench, Safer, & Levine, 2011; Peeters, Smith, Loewenstein, & Ubel, 2011; Van Boven & Loewenstein, 2005). For instance, widows and widowers have been found to underestimate the intensity of grief they will feel around the time of the anniversary of the death of a spouse (Carnelley et al., 2006).

The problem highlighted by our studies, then, is that use of the most common procedure for assessing forecasting accuracy makes it difficult to pin down how pronounced overestimation is, when it occurs, and why it occurs at some times and not others. In future research, it will be important to assess people's ability to predict the impact of events on their general emotional response, taking care that people clearly understand what it is they are being asked to predict.

Conclusions

To decide how strenuously to pursue or avoid outcomes, be it buying a car, changing careers, or having children, people attempt to predict how happy or unhappy those outcomes will make them. Overestimation in predicting emotion has received considerable attention from researchers and the media because it has important implications for models of decision making and for people's well-being (Gilbert & Wilson, 2007). Consistently overestimating future happiness would doom people to perpetual disappointment; consistently overestimating unhappiness, to perpetual dread (Buehler & McFarland, 2001). We found that people were highly accurate, however, when they both predicted and reported the intensity of their feelings about specific events. In contrast, people overestimated the impact of events on their general emotional well-being but this bias stemmed in part from use of a procedure in which people predict one thing but are later asked to report another. These findings demonstrate that people have more sophisticated self-knowledge than is commonly portrayed in the affective forecasting literature. Because early work on affective forecasting has led to the development of a burgeoning research field, we believe these findings are of considerable importance. Teasing apart people's ability to predict two distinct but concurrent features of emotional experience and identifying procedures that inflate the intensity bias represent theoretical and methodological refinements that enhance our understanding of people's affective forecasting abilities.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Andrade, E. B., & Van Boven, L. (2010). Feelings not foregone: Underestimating affective reactions to what does not happen. *Psychological Science, 21,* 706–711. doi:10.1177/0956797610368809

Ariely, D., Huber, J., & Wertenbroch, K. (2005). When do losses loom larger than gains? *Journal of Marketing Research, 42,* 134–138. doi:10.1509/jmkr.42.2.134.62283

Ariely, D., & Loewenstein, G. (2000). When does duration matter in judgment and decision making? *Journal of Experimental Psychology: General, 129,* 508–523. doi:10.1037/0096-3445.129.4.508

*Ayton, P., Pott, A., & Elwakili, N. (2007). Affective forecasting: Why can't people predict their emotions? *Thinking & Reasoning, 13,* 62–80. doi:10.1080/13546780600872726

Bacova, V., & Juskova, T. (2009). Risky choice and affective forecasting: Loss aversion in two culturally different student samples. *Studia Psychologica, 51,* 329–342.

Baron, J. (1992). The effect of normative beliefs on anticipated emotions. *Journal of Personality and Social Psychology, 63,* 320–330. doi:10.1037/0022-3514.63.2.320

Beedie, C. J., Terry, P. C., & Lane, A. M. (2005). Distinctions between emotion and mood. *Cognition & Emotion, 19,* 847–878. doi:10.1080/02699930541000057

*Böhm, G., & Pfister, H.-R. (2008). Anticipated and experienced emotions in environmental risk perception. *Judgment and Decision Making, 3,* 73–86.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). Comprehensive Meta-Analysis (Version 2) [Computer software]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester, England: Wiley. doi:10.1002/9780470743386

Bosson, J. K., Pinel, E. C., & Vandello, J. A. (2010). The emotional impact of ambivalent sexism: Forecasts versus real experiences. *Sex Roles, 62,* 520–531. doi:10.1007/s11199-009-9664-y

*Buehler, R., & McFarland, C. (2001). Intensity bias in affective forecasting: The role of temporal focus. *Personality and Social Psychology Bulletin, 27,* 1480–1493. doi:10.1177/01461672012711009

Buehler, R., McFarland, C., Spyropoulos, V., & Lam, K. C. H. (2007). Motivated prediction of future feelings: Effects of negative mood and mood orientation on affective forecasts. *Personality and Social Psychology Bulletin, 33,* 1265–1278.

*Carlsmith, K. M., Wilson, T. D., & Gilbert, D. T. (2008). The paradoxical consequences of revenge. *Journal of Personality and Social Psychology, 95,* 1316–1324. doi:10.1037/a0012165

Carnelley, K. B., Wortman, C. B., Bolger, N., & Burke, C. T. (2006). The time course of grief reactions to spousal loss: Evidence from a national probability sample. *Journal of Personality and Social Psychology, 91,* 476–492. doi:10.1037/0022-3514.91.3.476

Clore, G. L., & Huntsinger, J. R. (2009). How the object of affect guides its impact. *Emotion Review, 1,* 39–54. doi:10.1177/1754073908097185

Compton, R. J. (2003). The interface between emotion and attention: A review of evidence from psychology and neuroscience. *Behavioral and Cognitive Neuroscience Reviews, 2,* 115–129. doi:10.1177/1534582303002002003

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis.* New York, NY: Russell Sage Foundation.

Coughlan, R., & Connolly, T. (2001). Predicting affective responses to

unexpected outcomes. *Organizational Behavior and Human Decision Processes, 85,* 211–225. doi:10.1006/obhd.2000.2939

*Crawford, M. T., McConnell, A. R., Lewis, A. C., & Sherman, S. J. (2002). Reactance, compliance, and anticipated regret. *Journal of Experimental Social Psychology, 38,* 56–63. doi:10.1006/jesp.2001.1481

DeWall, C. N., & Baumeister, R. F. (2006). Alone but feeling no pain: Effects of social exclusion on physical pain tolerance and pain threshold, affective forecasting, and interpersonal empathy. *Journal of Personality and Social Psychology, 91,* 1–15. doi:10.1037/0022-3514.91.1.1

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95,* 542–575. doi:10.1037/0033-2909.95.3.542

Dillard, A. J., Fagerlin, A., Dal Cin, S., Zikmund-Fisher, B. J., & Ubel, P. A. (2010). Narratives that address affective forecasting errors reduce perceived barriers to colorectal cancer screening. *Social Science and Medicine, 71,* 45–52. doi:10.1016/j.socscimed.2010.02.038

Dohke, R., & Murata, K. (2009). Overestimation of anticipated and experienced regret after negative feedback. *Japanese Journal of Experimental Social Psychology, 48,* 150–158. doi:10.2130/jjesp.48.150

*Dunn, E. W., & Ashton-James, C. (2008). On emotional innumeracy: Predicted and actual affective responses to grand-scale tragedies. *Journal of Experimental Social Psychology, 44,* 692–698. doi:10.1016/j.jesp.2007.04.011

*Dunn, E. W., Biesanz, J. C., Human, L. J., & Finn, S. (2007). Misunderstanding the affective consequences of everyday social interactions: The hidden benefits of putting one's best face forward. *Journal of Personality and Social Psychology, 92,* 990–1005. doi:10.1037/0022-3514.92.6.990

Dunn, E. W., Brackett, M. A., Ashton-James, C., Schneiderman, E., & Salovey, P. (2007). On emotionally intelligent time travel: Individual differences in affective forecasting ability. *Personality and Social Psychology Bulletin, 33,* 85–93. doi:10.1177/0146167206294201

*Dunn, E. W., Wilson, T. D., & Gilbert, D. T. (2003). Location, location, location: The misprediction of satisfaction in housing lotteries. *Personality and Social Psychology Bulletin, 29,* 1421–1432. doi:10.1177/0146167203256867

Dunning, D., Heath, C., & Suls, J. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5,* 69–106. doi:10.1111/j.1529-1006.2004.00018.x

Duval, S., & Tweedie, R. L. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455–463. doi:10.1111/j.0006-341X.2000.00455.x

Eastwick, P. W., Finkel, E. J., Krishnamurti, T., & Loewenstein, G. (2008). Mispredicting distress following romantic breakup: Revealing the time course of the affective forecasting error. *Journal of Experimental Social Psychology, 44,* 800–807. doi:10.1016/j.jesp.2007.07.001

Elwyn, G., & Miron-Shatz, T. (2010). Deliberation before determination: The definition and evaluation of good decision making. *Health Expectations, 13,* 139–147. doi:10.1111/j.1369-7625.2009.00572.x

*Emanuel, A. S., Updegraff, J. A., Kalmbach, D. A., & Ciesla, J. A. (2010). The role of mindfulness facets in affective forecasting. *Personality and Individual Differences, 49,* 815–818. doi:10.1016/j.paid.2010.06.012

Falk, C., Dunn, E., & Norenzayan, A. (2010). Cultural variation in the importance of expected emotions in decision making. *Social Cognition, 28,* 609–629. doi:10.1521/soco.2010.28.5.609

*Fernandez-Duque, D., & Landers, J. (2008). "Feeling more regret than I would have imagined": Self-report and behavioral evidence. *Judgment and Decision Making, 3,* 449–456.

*Finkenauer, C., Gallucci, M., van Dijk, W. W., & Pollmann, M. (2007). Investigating the role of time in affective forecasting: Temporal influences on forecasting accuracy. *Personality and Social Psychology Bulletin, 33,* 1152–1166. doi:10.1177/0146167207303021

Frijda, N. H. (1994). Varieties of affect: Emotions and episodes, moods and sentiments. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 59–67). New York, NY: Oxford University Press.

Gaunt, R., Sindic, D., & Leyens, J. P. (2005). Intergroup relations in soccer finals: People's forecasts of the duration of emotional reactions of ingroup and outgroup soccer fans. *Journal of Social Psychology, 145,* 117–126. doi:10.3200/SOCP.145.2.117–126

Geers, A. L., & Lassiter, G. D. (1999). Affective expectations and information gain: Evidence for assimilation and contrast effects in affective experience. *Journal of Experimental Social Psychology, 35,* 394–413. doi:10.1006/jesp.1999.1377

Gertner, J. (2003, September 7). The futile pursuit of happiness. *The New York Times Magazine.* Retrieved from http://www.nytimes.com/2003/09/07/magazine/07HAPPINESS.html

Gilbert, D. T., & Ebert, J. E. J. (2002). Decisions and revisions: The affective forecasting of changeable outcomes. *Journal of Personality and Social Psychology, 82,* 503–513. doi:10.1037/0022-3514.82.4.503

Gilbert, D. T., Gill, M. J., & Wilson, T. D. (2002). The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes, 88,* 430–444. doi:10.1006/obhd.2001.2982

*Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science, 15,* 346–350. doi:10.1111/j.0956-7976.2004.00681.x

*Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 75,* 617–638. doi:10.1037/0022-3514.75.3.617

Gilbert, D. T., & Wilson, T. D. (2007, September 7). Prospection: Experiencing the future. *Science, 317,* 1351–1354. doi:10.1126/science.1144161

Gleitman, H., Reisberg, D., & Gross, J. (2010). *Psychology* (8th ed.). New York, NY: Norton.

Golub, S. A., Gilbert, D. T., & Wilson, T. D. (2009). Anticipating one's troubles: The costs and benefits of negative experiences. *Emotion, 9,* 277–281. doi:10.1037/a0014716

*Greitemeyer, T. (2009). The effect of anticipated affect on persistence and performance. *Personality and Social Psychology Bulletin, 35,* 172–186. doi:10.1177/0146167208326124

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, Vol. 3: Speech acts* (pp. 41–58). New York, NY: Academic Press.

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about the self and others. *Journal of Personality and Social Psychology, 56,* 1128–1139. doi:10.1037/0022-3514.59.6.1128

Hartnett, J., & Skowronski, J. J. (2008). Cash, money, woes: The match between a person's level of materialism and the materialistic (or non-materialistic) character of events alters affective forecasts. *North American Journal of Psychology, 10,* 647–664.

*Hartnett, J. L., & Skowronski, J. J. (2010). Affective forecasts and the Valentine's Day shootings at NIU: People are resilient, but unaware of it. *Journal of Positive Psychology, 5,* 275–280. doi:10.1080/17439760.2010.498615

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hoerger, M. (2012). Coping strategies and immune neglect in affective forecasting: Direct evidence and key moderators. *Judgment and Decision Making, 7,* 86–96.

*Hoerger, M., & Quirk, S. W. (2010). Affective forecasting and the Big Five. *Personality and Individual Differences, 49,* 972–976. doi:10.1016/j.paid.2010.08.007

*Hoerger, M., Quirk, S. W., Lucas, R. E., & Carr, T. H. (2009). Immune

neglect in affective forecasting. *Journal of Research in Personality, 43,* 91–94. doi:10.1016/j.jrp.2008.10.001

*Hoerger, M., Quirk, S. W., Lucas, R. E., & Carr, T. H. (2010). Cognitive determinants of affective forecasting errors. *Judgment and Decision Making, 5,* 365–373.

*Hsee, C. K., & Zhang, J. (2004). Distinction bias: Misprediction and mischoice due to joint evaluation. *Journal of Personality and Social Psychology, 86,* 680–695. doi:10.1037/0022-3514.86.5.680

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Newbury Park, CA: Sage.

Igou, E. R. (2004). Lay theories in affective forecasting: The progression of affect. *Journal of Experimental Social Psychology, 40,* 528–534. doi:10.1016/j.jesp.2003.09.004

Igou, R. (2008). "How long will I suffer?" versus "How long will you suffer?" A self–other effect in affective forecasting. *Journal of Personality and Social Psychology, 95,* 899–917. doi:10.1037/a0011619

Johnson, E. J., Steffel, M., & Goldstein, D. G. (2005). Making better decisions: From measuring to constructing preferences. *Health Psychology, 24,* S17–S22. doi:10.1037/0278-6133.24.4.S17

Kahneman, D., Ritov, I., & Schkade, D. (2000). Economic preferences or attitude expressions? An analysis of dollar responses to public issues. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 642–672). New York, NY: Cambridge University Press.

Kalisch, R., Wiech, K., Herrmann, K., & Dolan, R. J. (2006). Neural correlates of self-distraction from anxiety and a process model of cognitive emotion regulation. *Journal of Cognitive Neuroscience, 18,* 1266–1276. doi:10.1162/jocn.2006.18.8.1266

Kassam, K. S., Gilbert, D. T., Boston, A., & Wilson, T. D. (2008). Future anhedonia and time discounting. *Journal of Experimental Social Psychology, 44,* 1533–1537. doi:10.1016/j.jesp.2008.07.008

*Kawakami, K., Dunn, E., Karmali, F., & Dovidio, J. F. (2009, January 9). Mispredicting affective and behavioral responses to racism. *Science, 323,* 276–278. doi:10.1126/science.1164951

Keller, J., & Bless, H. (2009). Predicting future affective states: How ease of retrieval and faith in intuition moderate the impact of activated content. *European Journal of Social Psychology, 39,* 467–476. doi:10.1002/ejsp.586

*Kermer, D. A., Driver-Linn, E., Wilson, T. D., & Gilbert, D. T. (2006). Loss aversion is an affective forecasting error. *Psychological Science, 17,* 649–653. doi:10.1111/j.1467-9280.2006.01760.x

*Koo, M., Algoe, S. B., Wilson, T. D., & Gilbert, D. T. (2008). It's a wonderful life: Mentally subtracting positive events improves people's affective states, contrary to their affective forecasts. *Journal of Personality and Social Psychology, 95,* 1217–1224. doi:10.1037/a0013316

*Ku, G. (2008). Before escalation: Behavioral and affective forecasting in escalation of commitment. *Personality and Social Psychology Bulletin, 34,* 1477–1491. doi:10.1177/0146167208322559

*Kurtz, J. L., Wilson, T. D., & Gilbert, D. T. (2007). Quantity versus uncertainty: When winning one prize is better than winning two. *Journal of Experimental Social Psychology, 43,* 979–985. doi:10.1016/j.jesp.2006.10.020

*Lam, K. C. H., Buehler, R., McFarland, C., Ross, M., & Cheung, I. (2005). Cultural differences in affective forecasting: The role of focalism. *Personality and Social Psychology Bulletin, 31,* 1296–1309. doi:10.1177/0146167205274691

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin, 137,* 834–855. doi:10.1037/a0024244

Lench, H. C., Safer, M. A., & Levine, L. J. (2011). Focalism and the underestimation of future emotion: When it's worse than imagined. *Emotion, 11,* 278–285. doi:10.1037/a0022792

Levine, L. J. (1997). Reconstructing memory for emotions. *Journal of Experimental Psychology: General, 126,* 165–177. doi:10.1037/0096-3445.126.2.165

Levine, L. J., & Edelstein, R. S. (2009). Emotion and memory narrowing: A review and goal relevance approach. *Cognition & Emotion, 23,* 833–875. doi:10.1080/02699930902738863

Levine, L. J., & Kaplan, R. L. (2012). [Contributions of question misinterpretation and focalism to extreme affective forecasts]. Unpublished raw data.

Levine, L. J., Safer, M. A., & Lench, H. C. (2006). Remembering and misremembering emotions. In L. J. Sanna & E. C. Chang (Eds.), *Judgments over time: The interplay of thoughts, feelings, and behaviors* (pp. 271–290). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195177664.003.0015

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational and Human Decision Processes, 65,* 272–292. doi:10.1006/obhd.1996.0028

Loewenstein, G. (2001). A visceral account of addiction. In G. Loewenstein (Ed.), *Smoking: Risk, perception, and policy* (pp. 188–215). Thousand Oaks, CA: Sage.

Loewenstein, G. (2005). Hot–cold empathy gaps and medical decision making. *Health Psychology, 24,* S49–S56. doi:10.1037/0278-6133.24.4.S49

Loewenstein, G. (2007). Affect regulation and affective forecasting. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 180–203). New York, NY: Guilford Press.

Loewenstein, G. F., & Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review, 100,* 91–108. doi:10.1037/0033-295X.100.1.91

Loewenstein, G., & Schkade, D. (1999). Wouldn't it be nice? Predicting future feelings. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 85–105). New York, NY: Russell Sage Foundation.

MacInnis, D. J., & Patrick, V. M. (2006). Spotlight on affect: Affect and affective forecasting in impulse control. *Journal of Consumer Psychology, 16,* 224–231. doi:10.1207/s15327663jcp1603_4

*Mallett, R. K., Wilson, T. D., & Gilbert, D. T. (2008). Expect the unexpected: Failure to anticipate similarities leads to an intergroup forecasting error. *Journal of Personality and Social Psychology, 94,* 265–277. doi:10.1037/0022-3514.94.2.94.2.265

Marshall, M. A., & Brown, J. D. (2006). Emotional reactions to achievement outcomes: Is it really best to expect the worst? *Cognition and Emotion, 20,* 43–63. doi:10.1080/02699930500215116

Mathieu, M. T., & Gosling, S. D. (2012). The accuracy or inaccuracy of affective forecasts depends on how accuracy is indexed: A meta-analysis of past studies. *Psychological Science, 23,* 161–162. doi:10.1177/0956797611427044

Mellers, B. A., & McGraw, A. P. (2001). Anticipated emotions as guides to choices. *Current Directions in Psychological Science, 10,* 210–214. doi:10.1111/1467-8721.00151

*Meyvis, T., Ratner, R. K., & Levav, J. (2010). Why don't we learn to accurately forecast feelings? How misremembering our predictions blinds us to past forecasting errors. *Journal of Experimental Psychology: General, 139,* 579–589. doi:10.1037/a0020285

*Morewedge, C. K., Gilbert, D. T., Keysar, B., Berkovitz, M. J., & Wilson, T. D. (2007). Mispredicting the hedonic benefits of segregated gains. *Journal of Experimental Psychology: General, 136,* 700–709. doi:10.1037/0096-3445.136.4.700

Morewedge, C. K., Gilbert, D. T., & Wilson, T. D. (2005). The least likely of times: How remembering the past biases forecasts of the future. *Psychological Science, 16,* 626–630. doi:10.1111/j.1467-9280.2005.01585.x

*Nielsen, L., Knutson, B., & Carstensen, L. L. (2008). Affect dynamics, affective forecasting, and aging. *Emotion, 8,* 318–330. doi:10.1037/1528-3542.8.3.318

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics, 8,* 157–159. doi:10.2307/1164923

Osberg, T. M., & Shrauger, J. S. (1986). Self-prediction: Exploring the parameters of accuracy. *Journal of Personality and Social Psychology, 51,* 1044–1057. doi:10.1037/0022-3514.51.5.1044

Patrick, V. M., Chun, H. H., & MacInnis, D. J. (2009). Affective forecasting and self-control: Why anticipating pride wins over anticipating shame in a self-regulation context. *Journal of Consumer Psychology, 19,* 537–545. doi:10.1016/j.jcps.2009.05.006

Peeters, Y., Smith, D. M., Loewenstein, G., & Ubel, P. A. (2011). After adversity strikes: Predictions, recollections and reality among people experiencing the onset of adverse circumstances. *Journal of Happiness Studies.* Advance online publication. doi:10.1007/s10902-011-9281-7

*Pollmann, M. H., & Finkenauer, C. (2009). Empathic forecasting: How do we predict other people's feelings? *Cognition & Emotion, 23,* 978–1001. doi:10.1080/02699930802264895

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879–891. doi:10.3758/BRM.40.3.879

Rachman, S., & Bichard, S. (1988). The overprediction of fear. *Clinical Psychology Review, 8,* 303–312. doi:10.1016/0272-7358(88)90093-1

Raudenbush, S. W. (1994). Random effects model. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–321). New York, NY: Russell Sage Foundation.

*Riis, J., Loewenstein, G., Baron, J., Jepson, C., Fagerlin, A., & Ubel, P. A. (2005). Ignorance of hedonic adaptation to hemodialysis: A study using ecological momentary assessment. *Journal of Experimental Psychology: General, 134,* 3–9. doi:10.1037/0096-3445.134.1.3

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

Safer, M. A., Christianson, S. A., Autry, M. W., & Osterlund, K. (1998). Tunnel memory for traumatic events. *Applied Cognitive Psychology, 12,* 99–117. doi:10.1002/(SICI)1099-0720(199804)12:2<99::AID-ACP509>3.0.CO;2-7

Samanez-Larkin, G. R., Gibbs, S. E. B., Khanna, K., Nielsen, L., Carstensen, L. L., & Knutson, B. (2007). Anticipation of monetary gain but not loss in healthy older adults. *Nature Neuroscience, 10,* 787–791. doi:10.1038/nn1894

Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality, 61,* 317–342. doi:10.1111/j.1467-6494.1993.tb00283.x

*Sanna, L. J., & Schwarz, N. (2004). Integrating temporal biases: The interplay of focal thoughts and accessibility experiences. *Psychological Science, 15,* 474–481. doi:10.1111/j.0956-7976.2004.00704.x

Schkade, D. A., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science, 9,* 340–346. doi:10.1111/1467-9280.00066

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54,* 93–105. doi:10.1037/0003-066X.54.2.93

Seta, J. J., Haire, A., & Seta, C. E. (2008). Averaging and summation: Positivity and choice as a function of the number and affective intensity of life events. *Journal of Experimental Social Psychology, 44,* 173–186. doi:10.1016/j.jesp.2007.03.003

*Sevdalis, N., & Harvey, N. (2007). Biased forecasting of postdecisional affect. *Psychological Science, 18,* 678–681. doi:10.1111/j.1467-9280.2007.01958.x

*Sevdalis, N., Harvey, N., & Bell, A. (2009). Affective equilibria in the endowment effect. *Journal of Behavioral Finance, 10,* 89–100. doi:10.1080/15427560902720545

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper, & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–281). New York, NY: Russell Sage Foundation.

Sheldon, K. M., Gunz, A., Nichols, C. P., & Ferguson, Y. (2010). Extrinsic value orientation and affective forecasting: Overestimating the rewards, underestimating the costs. *Journal of Personality, 78,* 149–178. doi:10.1111/j.1467-6494.2009.00612.x

*Sieff, E. M., Dawes, R. M., & Loewenstein, G. (1999). Anticipated versus actual reaction to HIV test results. *American Journal of Psychology, 112,* 297–311. doi:10.2307/1423355

*Smith, D., Loewenstein, G., Jepson, C., Jankovich, A., Feldman, H., & Ubel, P. (2008). Mispredicting and misremembering: Patients with renal failure overestimate improvements in quality of life after a kidney transplant. *Health Psychology, 27,* 653–658. doi:10.1037/a0012647

Smith, E. R., & Mackie, D. M. (2009). Surprising emotions. *Science, 323,* 215–216. doi:10.1126/science.1168650

Solomon, R. C. (1973). Emotions and choice. *Review of Metaphysics, 27,* 20–41.

Strack, F., & Schwarz, N. (2007). Asking questions: Measurement in the social sciences. In M. Ash & T. Sturm (Eds.), *Psychology's territories: Historical and contemporary perspectives from different disciplines* (pp. 225–250). Mahwah, NJ: Erlbaum.

Sweeny, K., & Shepperd, J. A. (2010). The costs of optimism and the benefits of pessimism. *Emotion, 10,* 750–753. doi:10.1037/a0019016

*Tomlinson, J. M., Carmichael, C. L., Reis, H. T., & Aron, A. (2010). Affective forecasting and individual differences: Accuracy for relational events and anxious attachment. *Emotion, 10,* 447–453. doi:10.1037/a0018701

Totterdell, P., Parkinson, B., Briner, R. B., & Reynolds, S. (1997). Forecasting feelings: The accuracy and effects of self-predictions of mood. *Journal of Social Behavior and Personality, 12,* 631–650.

Ubel, P. A., Loewenstein, G., Hershey, J., Baron, J., Mohr, T., Asch, D., & Jepson, C. (2001). Do nonpatients underestimate the quality of life associated with chronic health conditions because of a focusing illusion? *Medical Decision Making, 21,* 190–199.

Van Boven, L., & Ashworth, L. (2007). Looking forward, looking back: Anticipation is more evocative than retrospection. *Journal of Experimental Psychology: General, 136,* 289–300. doi:10.1037/0096-3445.136.2.289

Van Boven, L., & Loewenstein, D. (2005). Empathy gaps in emotional perspective taking. In S. Hodges & B. Malle (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 284–297). New York, NY: Guilford Press.

Van Boven, L., Loewenstein, G., & Dunning, D. (2005). The illusion of courage in social predictions: Underestimating the impact of fear of embarrassment on other people. *Organizational Behavior and Human Decision Processes, 96,* 130–141. doi:10.1016/j.obhdp.2004.12.001

*van Dijk, W. W. (2009). How do you feel? Affective forecasting and the impact bias in track athletics. *Journal of Social Psychology, 149,* 343–348. doi:10.3200/SOCP.149.3.343-348

*van Dijk, W. W., Finkenauer, C., & Pollmann, M. (2008). The misprediction of emotions in track athletics: Is experience the teacher of all things? *Basic and Applied Social Psychology, 30,* 369–376. doi:10.1080/01973530802502358

Walsh, E., & Ayton, P. (2009). My imagination versus your feelings: Can personal affective forecasts be improved by knowing other peoples' emotions? *Journal of Experimental Psychology: Applied, 15,* 351–360. doi:10.1037/a0017984

Welsch, H., & Kühling, J. (2010). Pro-environmental behavior and rational consumer choice: Evidence from surveys of life satisfaction. *Journal of Economic Psychology, 31,* 405–420. doi:10.1016/j.joep.2010.01.009

Wesp, R., Sandry, J., Prisco, A., & Sarte, P. (2009). Affective forecasts of future positive events are tempered by consideration of details. *American Journal of Psychology, 122,* 167–174.

Whalen, C. K., Henker, B., King, P. S., Jamner, L. D., & Levine, L. (2004). Adolescents react to the events of September 11, 2001: Focused versus

ambient impact. *Journal of Abnormal Child Psychology, 32,* 1–11. doi:10.1023/B:JACP.0000007576.94501.a0

*Wilson, T. D., Centerbar, D. B., Kermer, D. A., & Gilbert, D. T. (2005). The pleasures of uncertainty: Prolonging positive moods in ways people do not anticipate. *Journal of Personality and Social Psychology, 88,* 5–21. doi:10.1037/0022-3514.88.1.5

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 345–411). New York, NY: Elsevier.

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science, 14,* 131–134.

Wilson, T. D., Meyers, J., & Gilbert, D. T. (1999). [The durability bias in prediction and postdiction]. Unpublished raw data.

*Wilson, T. D., Meyers, J., & Gilbert, D. T. (2003). "How happy was I, anyway?" A retrospective impact bias. *Social Cognition, 21,* 421–446. doi:10.1521/soco.21.6.421.28688

Wilson, T. D., Wheatley, T. P., Kurtz, J. L., Dunn, E. W., & Gilbert, D. T. (2004). When to fire: Anticipatory versus postevent reconstrual of uncontrollable events. *Personality and Social Psychology Bulletin, 30,* 340–351. doi:10.1177/0146167203256974

*Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 78,* 821–836. doi:10.1037/0022-3514.78.5.821

Winter, L., Moss, M. S., & Hoffman, C. (2009). Affective forecasting and advance care planning: Anticipating quality of life in future health statuses. *Journal of Health Psychology, 14,* 447–456. doi:10.1177/1359105309102201

*Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science, 14,* 520–524. doi:10.1111/1467-9280.03455

Wood, S. L., & Bettman, J. R. (2007). Predicting happiness: How normative feeling rules influence (and even reverse) durability bias. *Journal of Consumer Psychology, 17,* 188–201. doi:10.1016/S1057-7408(07)70028-1

Zaki, J., & Ochsner, K. (2011). Reintegrating the study of accuracy into social cognition research. *Psychological Inquiry, 22,* 159–182. doi:10.1080/1047840X.2011.551743

Zeelenberg, M., Nelissen, R. M. A., Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision Making, 3,* 18–27.

Zeelenberg, M., van Dijk, W. W., Manstead, A. S. R., & van der Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition & Emotion, 14,* 521–541. doi:10.1080/026999300402781

Zhao, S., & Meyer, R. J. (2007). Biases in predicting preferences for the whole visual patterns from product fragments. *Journal of Consumer Psychology, 17,* 292–304. doi:10.1016/S1057-7408(07)70039-6

Received September 19, 2011
Revision received May 14, 2012
Accepted May 21, 2012 ■