American Psychological Association

Journal of Personality and Social Psychology

Manuscript version of

When and Why People Misestimate Future Feelings: Identifying Strengths and Weaknesses in Affective Forecasting

Heather C. Lench, Linda J. Levine, Kenneth Perez, Zari Koelbel Carpenter, Steven J. Carlson, Shane W. Bench, Yidou Wan

Funded by:

- National Science Foundation

Abstract

People try to make decisions that will improve their lives and make them happy, and to do so, they rely on affective forecasts - predictions about how future outcomes will make them feel. Decades of research suggest that people are poor at predicting how they will feel and that they commonly overestimate the impact that future events will have on their emotions. Recent work reveals considerable variability in forecasting accuracy. This investigation tested a model of affective forecasting that captures this variability in bias by differentiating emotional intensity, emotional frequency, and mood. Two field studies examined affective forecasting in college students receiving grades on a midterm exam (Study 1, $N = 643$), and U.S. citizens after the outcome of the 2016 presidential election (Study 2, $N = 706$). Consistent with the proposed model, participants were more accurate in forecasting the intensity of their emotion and less accurate in forecasting emotion frequency and mood. Overestimation of the effect of the event on mood increased over time since the event. Three experimental studies examined mechanisms that contribute to differential forecasting accuracy. Biases in forecasting intensity were caused by changes in perceived event importance; biases in forecasting frequency of emotion were caused by changes in the frequency of thinking about the event. This is the first direct evidence mapping out strengths and weaknesses for different types of affective forecasts and the factors that contribute to this pattern.

KEYWORDS:
Affective forecasting
Emotion
Prediction
Decision making
Bias

When and Why People Misestimate Future Feelings:

Identifying Strengths and Weaknesses in Affective Forecasting

People try to make decisions that will improve their lives and make them happy, and to do so they rely on affective forecasts - guesses about how future events will make them feel. They direct effort and resources toward attaining outcomes that will result in happiness and avoiding outcomes that will induce misery. The stronger the emotional reactions that people expect to have, the harder they will work to attain or avoid outcomes (Mellers & McGraw, 2001; Morewedge & Buechel, 2013). Because of this, inaccuracy in affective forecasting has been identified as one of the leading barriers to effective decision making and well-being (Loewenstein, 2007). Life is much harder when people cannot foresee what lies ahead, and people's inability to accurately predict the emotional consequences of their choices is thought to be a primary source of poor decisions and unhappiness. Given the importance of affective forecasts to choice, a precise understanding of how people forecast their future emotions is needed in order to improve decision making.

An Evaluation of Impact Bias

Multiple early studies on affective forecasting demonstrated that people are poor at predicting how strongly they will feel emotions in the future. These studies revealed that people have a persistent tendency to overestimate the strength of their reactions to a variety of events, such as enjoyment of holidays, being denied tenure, and the victory of a favored sports team (e.g., Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 2002; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000). Further investigations revealed that this overestimation results in part because people often fail to anticipate how quickly they will adapt to events (adaptation neglect) and they often fail to consider other events that will occupy their thoughts and impact their emotions in the future (focalism; Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Wilson & Gilbert, 2008). The persistent tendency to exaggerate future emotional reactions is referred to as the “impact bias” (e.g., Gilbert et al., 2002), and includes overestimates of the intensity (how strong a reaction will be) and duration (how long a reaction will last) of future emotion. The robustness and pervasiveness of this bias has been widely accepted in the field, and is the subject of popular books (Gilbert, 2006), disseminated in the media, and reported in introductory psychology textbooks.

The simplicity of the idea that people have a general tendency to overestimate the impact events will have on their emotions is intuitively attractive. However, accumulated evidence suggests that this idea is too simple to account for research findings and that a more precise theoretical explanation is required to move the field forward. Studies have documented conditions that result in underestimation in forecasting future emotion. Underestimation occurs when evocative characteristics of events, which were not salient at the time of forecasting, become salient during the emotional experience. For example, people underestimate how much they will enjoy social interactions with a stranger which entail presenting themselves positively (Dunn, Biesanz, Human, & Finn, 2007). People underestimate their reactions to experiences that include unforseen visceral cues which intensity emotion (e.g., the smell of cookies baking; Lench, Safer, & Levine, 2011). More generally, forecasters tend to focus on basic characteristics of a future event that are typically diagnostic of its hedonic impact whereas experiencers are immersed in the more richly-nuanced event itself (Buechel, Zhang, & Morewedge, 2017; Ebert & Meyvis, 2014). Thus, a mismatch in the information available at the time of forecasting versus experiencing emotion can lead to underestimating, as well as overestimating, future emotion.

Beyond the presence of mixed findings regarding the direction of bias, findings from an additional line of research cast doubt on whether the impact bias is pervasive and suggest that the degree of bias varies depending on what feature of emotion people are asked to forecast. People are far more accurate when they are asked to forecast the intensity of their emotional response to specific events compared to forecasting their general emotion or mood (Doré, Meksin, Mather, Hirst, & Ochsner, 2016; Kaplan, Levine, Lench, & Safer, 2016; Levine, Lench, Kaplan, & Safer, 2012, 2013). In addition, a series of studies demonstrated that procedures commonly used to assess forecasting accuracy inflate the magnitude of the impact bias because people misunderstand the questions asked by researchers (Levine et al., 2012). Inaccuracy occurs because people often believe that they have been asked to predict one feature of emotion whereas experimenters later assess another. The procedure in question is this: Most affective forecasting studies ask people to imagine that an event has occurred (e.g., imagine that you are denied tenure) and to rate how they will feel in general after a specified period of time (e.g., a year later, in general, how happy will you feel?; Gilbert et al., 1998). After the event occurs and the specified period of time has elapsed, participants rate their current feelings without reference to the target event (e.g., “In general, how happy would you say you are these days?”). Levine and colleagues (2012) demonstrated that, in the context of having just been asked to imagine a specific future event, the majority of participants interpreted the request to predict how they would feel *in general* as asking them to predict the *intensity of emotion they would feel when thinking about that event*. This misinterpretation virtually guarantees overestimation in forecasting emotion.

Further, a series of studies demonstrated that use of alternative procedures reduced or eliminated the impact bias, and revealed that participants were very accurate when forecasting the intensity of future emotion (Levine et al., 2012). A meta-analytic review of the affective forecasting literature showed that participants only reliably overestimated their future emotions when they were asked the "general" question with a delay after the event (Levine et al., 2012). Forecasts were far more accurate when participants were asked to forecast the intensity of their response to the event or asked to report their general emotion immediately after the event occurred (i.e., when they were still thinking about the event). The results of these studies identify procedures that inflate bias in forecasts, and also document that the feature of emotion being forecast partially explains the mixed findings in the forecasting literature. Most importantly, these findings suggest that any theory of affective forecasting biases must take into account the particular feature of emotional experience that people are attempting to predict.

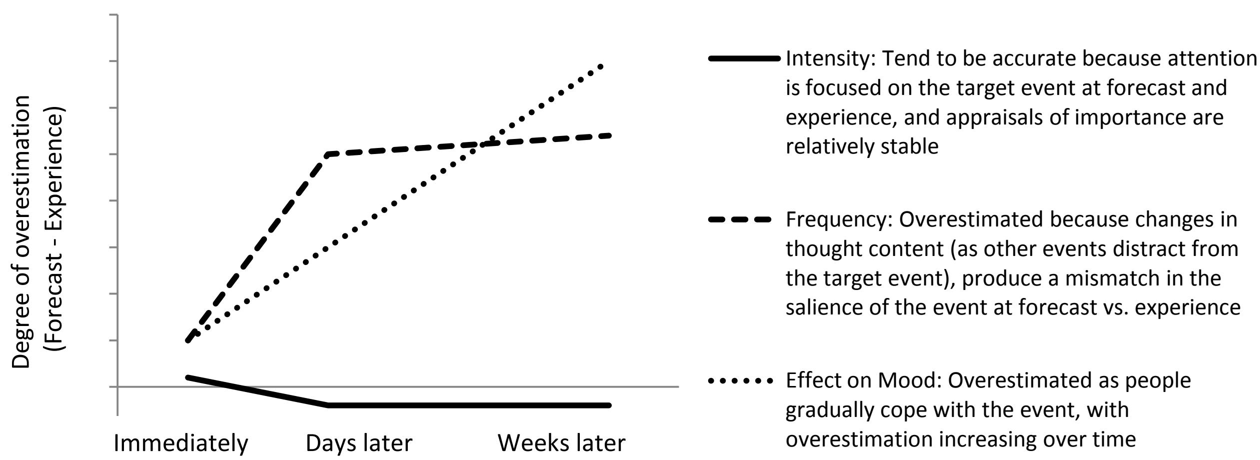
The "impact bias" subsumes forecasts of the intensity of future feeling, the frequency of future feeling, and the effect of an event on mood, without distinguishing among these features of emotional experience. Yet the findings noted above suggest that people are better at predicting some of these features of emotion than others. Below, we present a new theoretical model of the patterns of biases that result when people forecast different features of emotion. According to our model, people show less bias when forecasting the intensity of emotion they will experience about an event in the future. However, they overestimate the frequency of their emotional response to the event after it occurs, as well as the effect of the event on mood. Each of these features is potentially critical for optimizing decision making. To date, however, there has been no systematic investigation of people's ability to predict distinct features of their emotional experience. The accumulation of paradoxical findings in the affective forecasting literature cries out for an investigation with this level of specificity to explain when and why people forecast accurately or show an impact bias. In summary, further investigation is imperative because of the central role affective forecasting plays in decisions large and small, the disparities in the empirical evidence for an impact bias, and the recent controversy over research methods.

The Differentiated Model of Affective Forecasting

We propose that people have a combination of forecasting strengths and weaknesses that allow them to predict certain features of emotion far more accurately than others. We further propose that the extent to which forecasting accuracy changes over time depends on the feature of emotion people are predicting. The theoretical construct of "impact bias" does not differentiate people's abilities to forecast different features of emotion, such as the intensity versus frequency of emotions about an event. The initial assumption was that forecasting bias and the mechanisms contributing to bias would be similar across features of emotion (Wilson & Gilbert, 2003), but the recent evidence reviewed above demonstrates that this assumption is incorrect (Levine et al., 2012). We identified three features of emotion from methods employed in the forecasting literature that have also been established as relevant to emotion experience (e.g., Kuppens & Verduyn, 2017), and examined forecasts for these features of emotion: the *intensity* of emotion about an event, the *frequency* of emotion about an event, and the *effect of an event on mood*. Below we describe the patterns of biases hypothesized, based on characteristics of each feature of emotion. These predictions focus on absolute accuracy (i.e., the difference between forecast and experienced emotion), and not relative accuracy (i.e., the degree to which individuals can predict their response relative to the response of others). This choice was made because absolute accuracy is the focus of the majority of studies and theory related to affective forecasting, and because the identified processes are specific to absolute accuracy. We return to this issue in the discussion.

Levine et al. (2012) provided initial evidence that the degree of bias depends on the feature of emotion participants are predicting. Forecasting questions that made it clear that participants were being asked to predict their overall mood, rather than the intensity of their feelings specifically about the target event, reduced overestimation but did not eliminate it. Questions that targeted emotional intensity (i.e., participants were asked both to predict and to report the intensity of their feelings about an event) led to strikingly accurate forecasts (and even a tendency toward underestimation). The aims of the present investigation were to assess people's ability to forecast key features of emotion and to identify mechanisms contributing to biases in these forecasts. Our hypotheses for emotion intensity, emotion frequency, and effect on mood are shown in Figure 1.

Figure 1. Hypotheses derived from a differentiated model of affective forecasting.



We hypothesized that people can accurately forecast the *intensity* of future emotion about an event, and that forecasts of intensity will be less biased than forecasts of other features of emotion. This hypothesis is based on well-documented accounts of 1) the effects of emotional responses on attention, and 2) the causes of emotional intensity. First, intense emotion, particularly when negative, narrows the focus of attention to central features of events at the expense of peripheral features (for reviews see Compton et al., 2003; Lench & Bench, 2015; Levine & Edelstein, 2009). When forecasting their future feelings, people also tend to focus on an event's salient features (Gilbert et al., 1998). In the case of forecast intensity, the combination of these tendencies means that the features of emotion-eliciting events that are salient at the time of forecasting are also likely to be salient when people are later thinking about those events. In other words, the narrow focus of attention on the target event at both forecast and experience should promote accurate forecasts of emotional intensity. Second, we theorized that forecasts of intensity are based on appraisals that tend to be stable over time. Appraisals are evaluations of events that reflect an event's relation to an individual's goals. Past studies have demonstrated that the intensity of experienced emotion about an event is tightly coupled with the perceived importance of that event (Verduyn, Delvaux, Van Coillie, Tuerlinckx, & Van Mechelen, 2009; Verduyn, Van Mechelen, & Frederix, 2012; Verduyn, Van Mechelen, Tuerlinckx, & Scherer, 2013). Contributing to accuracy in forecasting the intensity of emotion, people's perceptions of the importance of major events tends to be stable over time (McAdams & Olson, 2010). Consistent with this hypothesis, studies have revealed that people's forecasts tend to be accurate both when they forecast the intensity of emotion they will experience shortly after an event occurs, and when they forecast the intensity of emotion they will experience when an event comes to mind after a delay (Crawford, McConnell, Lewis, & Sherman, 2002; Doré et al., 2016; Lench et al., 2011; Levine et al., 2012).

We hypothesized greater inaccuracy, and a tendency toward overestimation, for forecasts of the *frequency* of emotional reactions to events. This hypothesis is based on extensive evidence that people tend to focus on the target event at the time of forecast, but that once the event occurs attention is likely to be captured by other ongoing events, resulting in overestimation of the event's emotional impact (e.g., Gilbert et al., 1998; Wilson et al., 2000). Previous accounts have posited that this mechanism is relevant to forecasts of any emotion feature. In contrast, we theorized that this mechanism – overestimating how often thoughts will be focused on the event - is particularly relevant to people's estimates of how frequently they will experience emotional reactions to the event in the future. Although the emotion-eliciting event is salient at the time of forecast, the contents of a person's thoughts change rapidly and unpredictably after an event occurs depending on what else is happening at the time. Unlike forecasts of intensity, then, the basis of forecasted emotional frequency is unstable over time because people are unlikely to be able to predict the content of their future thoughts. This should result in biased predictions that overestimate the frequency of emotion.

In the affective forecasting literature, the emotional impact of an event is often conceptualized as the effect of the event on a person's overall mood (e.g., Gilbert et al., 1998), but participants are rarely asked explicitly to forecast how much events will affect their mood (Levine et al., 2012). Without direct evidence regarding forecasting accuracy for the effect of events on mood, we hypothesized that forecasts of the effect of an event on *overall mood* would be more biased than intensity forecasts, and would tend toward overestimation. This hypothesis is based on evidence that people adjust to and cope with events after they occur, with mood returning to baseline levels gradually after positive and negative events (Diener, Suh, Lucas, & Smith, 1999; Headey & Wearing, 1989; Lucas & Donnellan, 2007). Even the impact of major life events, such as serious disability or large lottery wins, decreases over time (Brickman, Coates, & Janoff-Bulman, 1978; Suh, Diener, & Fujita, 1996). Studies demonstrated that "impact bias" can result from people neglecting to account for their ability to adjust to and cope with events (adaptation neglect, Gilbert et al., 1998; Wilson & Gilbert, 2008). Based on past evidence that general mood returns to baseline levels over time after an event, we theorized that this mechanism – failure to anticipate that one will adapt to an event in the future - was particularly relevant to forecasts of the effect of an event on mood. Following from this theorized relationship, we hypothesized that bias in forecasts the effect of an event on mood should increase over time since the event, as people continue to cope with and adjust to the event. Given past work linking surprise to the duration of an emotional experience (e.g., Verduyn et al., 2009), it was possible that the rate of accommodation to events depends on how much the event violates prior expectations (i.e., "surprisingness"). We therefore also examined surprise as a potential mechanism that contributes to bias in forecasting the effect of events on mood.

In summary, we theorize that people have unique strengths and weakness in forecasting different features of emotion. The degree of bias, and the mechanisms contributing to bias, were expected to differ for forecasts of emotional intensity, emotional frequency, and effect of an event on mood.

The Present Investigation

We examined two primary questions in the present investigation: 1) does forecasting bias vary depending on the feature of emotion people are forecasting?, and 2) what mechanisms account for biases in forecasting different features of emotion? To address these two questions, we used a combination of field and laboratory studies. General approach and hypotheses were posted before data collection and analysis at https://osf.io/ph4xf/, as well as in the grant proposal that supported data collection.

Field studies were used to address the first question, by examining patterns of forecasting biases in real-world settings with consequential events that elicited strong emotional responses. These included college students receiving their grades on midterm exams in introductory courses and United States citizens responding to the outcome of the 2016 presidential election. These events were chosen because past studies have revealed an impact bias in forecasts for these events (e.g., Levine et al., 2012; Wilson, Meyers, & Gilbert, 2003) and because they occurred on a specified and predetermined date. Participants reported their forecast and experienced intensity of emotion about the event, the frequency of emotion, and the effect of the event on mood. Participants also reported their forecast and experienced emotion "in general," using the question format that is common in affective forecasting studies. We included this question only as a reference point for the broader literature, and recommend caution interpreting these findings as past studies have demonstrated that participants frequently misunderstand these questions (Levine et al., 2012).

The laboratory studies were used to examine the second question, regarding the mechanisms that contribute to biases in forecasts of different features of emotion. Although field studies can be informative about relationships among variables, they do not permit the type of experimental control that is necessary for examining causal relationships. To this end, the laboratory studies were designed to systematically test the effects of three cognitive features on bias in forecasting different features of emotion: perceived importance of the event, frequency of thinking about the event, and surprisingness of the event. As described above, we hypothesized that perceived importance would be particularly relevant to forecasting bias for intensity of emotion. Although perceived importance tends to be stable over time, we hypothesized that people would be inaccurate in their forecasts about the intensity of their future emotion to the extent that their appraisal of the event's importance changed. In contrast to prior conceptualizations in the forecasting literature that posit changes in frequency of thought (i.e., focalism) as a mechanism contributing to all forecasting biases, we hypothesized that changes in frequency of thinking about an event would be particularly relevant to bias in forecasts of the frequency of emotion about an event (Tully & Meyvis, 2017). Based on recent work demonstrating the contribution of surprise to the duration of emotional responses (e.g., Verduyn et al., 2009), we hypothesized that the suprisingness of the event would be particularly relevant to bias in forecasts of the effect of the event on mood. Importantly, because we experimentally manipulated perceptions of events in the laboratory studies, these studies address mechanisms that contribute to biases in forecasts of different features of emotion but these studies cannot be used to draw conclusions about the relative accuracy of forecasts of different features of emotion (an issue that is addressed in Studies 1 and 2).

The present investigation is the first to systematically differentiate among biases in forecasts of the intensity of emotion, frequency of emotion, and effect of events on mood. Based on a new theoretical model and preliminary findings, people are hypothesized to show less bias in forecasting the intensity of their emotion, but more bias and overestimation when forecasting the frequency of emotion and the effect of events on mood. This investigation also unpacks the differing mechanisms that result in bias in forecasting these three features of emotion. Addressing these questions is critically important, not only for understanding people's ability to forecast their emotions, but also for understanding how forecasting biases relate to the broader literature on motivation and emotion.

Study 1: Exam Grades in an Introductory University Course

In Study 1, university students forecast and reported their emotions concerning their first midterm exam in an introductory course. The purpose of this study was to examine patterns of forecasting bias for different features of emotion. We also included questions regarding the cognitive features of interest (importance, frequency of thinking, surprise) as a preliminary examination of the relevance of these features to forecasting biases.

Methods

Participants. Data collection was planned for approximately 500 participants, based on estimates that this would provide sufficient power to detect a small effect size. (An a priori g*power computation for the difference between two dependent means gave a total sample size of 327 to have power of .95 to detect an effect size of .20; an a priori g*power computation for a MANOVA with 4 repeated measures gave a total sample size of 216 to have power of .95 to detect an effect size *f* of .10). This estimated effect size would be conservatively small, given that a recent meta-analysis of the affective forecasting literature (Levine et al., 2012) revealed an overall average effect size (*g*) of .55, with a 95% confidence interval from .42 to .68. Recruitment was completed through emails to the roster of undergraduate courses, and therefore it was not possible to perfectly target a sample size, and we intentionally oversampled due to anticipated non-respondents, as well as attrition over time, for a total of 707 participants who completed the first survey. Participants were recruited from undergraduate introductory psychology courses and received course credit for completing the surveys. Data was collected online, and the study was conducted at two universities that varied in the demographic profile of their students.

Participants were excluded from analyses if they did not complete the survey on which they reported their emotional experience after they found out their exam grade, or did not complete the question that was used to categorize the valence of the event (by identifying whether the grade they received was worse than, better than, or the same as expected). The final sample included 643 participants, with 359 participants enrolled in a large public California university and 284 participants enrolled in a large public Texas university. This recruitment provided a post hoc power of .9990 to detect a .20 *d* effect size. The sample included 77% women, with an average age of 19.56 years (*SD* = 2.73).

Procedures and materials. This study was part of a larger investigation, and only methods and procedures relevant to the present hypotheses are reported here. Two weeks before their first exam in an introductory psychology course, participants forecast how they would feel two days after they found out their grade on the exam. Two days after finding out their grade, participants reported their experienced reactions.

Forecasts (Time 1). Participants were asked what grade they expected to receive on an upcoming exam. They then forecast how they would feel if they received a higher grade than expected, a lower grade than expected, and the grade they expected. Participants rated happiness and unhappiness for three features of their emotional experience: intensity, frequency, and effect on mood.

Specifically, participants were prompted to “Suppose you get a grade that is [higher than/lower than/the same as] you expect” and then, using a question format similar to past affective forecasting studies, participants were asked, “Two days after you find out your grade, how will you feel in general?” They forecast how happy and how unhappy they would feel in general on scales ranging from *not at all* (1) to *extremely* (9). As noted in the introduction, this question format has been demonstrated to inflate bias due to a methodological artifact (Levine et al., 2012), in that participants often misunderstand the question they are being asked. We included this question because it is the most common format in the associated literature and is therefore a useful reference point for the degree to which bias is present in forecasts for the three features of emotion relative to the most commonly used question. However, because participants misinterpret this question, we recommend limited inferences from this data. We refer to this item as “ambiguous emotion” in results.

Participants then forecast their feelings using questions that focused specifically on their response to the event, including the intensity of their emotion, the frequency of their emotion, and the effect of the event on mood. For intensity, they were asked: “Two days after you find out your grade, how will you feel about receiving that grade? Happy/Unhappy” on scales ranging from *not at all* (1) to *extremely* (9). They also forecast the frequency of their emotion, “Suppose it is two days after you find out your grade, and you get a grade that is higher than you expect. Tell us how you will answer the following question at that time: Yesterday, how frequently did you feel this way about getting an exam grade that was [higher than/lower than/the same as] you expected. Happy/Unhappy” on scales from *not at all* (1) to *constantly* (9). The time frame was set as “yesterday” in order to ensure that all participants were referencing a time period with the same number of hours, regardless of the specific time that they completed the survey concerning their emotional experience. Participants forecast the effect of the event on their overall mood, “Suppose you get a grade that is [higher than/lower than/the same as] you expect. Two days after you find out your grade, how much of an impact will getting that grade have on your overall mood?” on a scale ranging from *no impact at all* (1) to *extremely impactful* (9).

The primary question addressed by Study 1 was the degree of bias in forecasts of different features of emotion, but questions were also included to preliminarily assess the relevance of three cognitive features theorized to contribute to biases. Participants forecast the frequency of thinking about the exam outcome, “Suppose it is two days after you find out your grade, and you get a grade that is [higher than/lower than/the same as] you expect. Tell us how you will answer the following question at that time: Yesterday, how frequently did you think about your exam grade?” using the same scale as for frequency of feeling. Again, the time period “yesterday” was specified for frequency measurements to keep the number of hours consistent across participants. Participants also reported their appraisals of the exam grade, including importance, “How important is your exam grade to you?”; and surprise, “Suppose you get a grade that is [higher than/lower than/the same as] than you expected. How surprised will you be by your exam grade?” Participants rated these appraisals on a scale from *not at all* (1) to *extremely* (9).

Experience (Time 2). Participants completed the second survey between 7 a.m. and midnight of the day that fell two days after their exam grades were released. As a pair to the “ambiguous emotion” question asked at forecast, consistent with the typical affective forecasting procedure, participants first reported their current feelings with no reference to the exam, “How are you currently feeling in general? Happy/Unhappy” on scales ranging from *not at all* (1) to *extremely* (9).

Participants verified that they had seen their exam grade, and reported whether the grade they received was worse than they expected, the same as they expected, or better than they expected. They then reported their experienced emotional intensity, emotional frequency, and the effect of the exam grade outcome on mood. These measurements used the same questions and scales as at forecast, with verbs changed from future to present tense. Participants also reported their current appraisals of the importance of the exam grade, the frequency with which they were thinking about the event, as well as their degree of surprise at their exam grade, using the same questions as at forecast with verb tenses changed.

Results

Preliminary analyses. A total of 188 participants (29.2%) received a grade that was higher than expected and an additional 128 (19.9%) received the grade they expected. Preliminary analyses suggested that happiness was the primary anticipated response both for getting a grade that was higher than expected (M_{happy} = 7.69 vs. $M_{unhappy}$ = 1.24) and for getting a grade that was the same as expected (M_{happy} = 6.83 vs. $M_{unhappy}$ = 1.77). Because patterns were similar for these two groups, they were combined. A total of 327 (50.9%) received a grade that was lower than they had expected, and unhappiness was the primary anticipated response for this group (M_{happy} = 2.68 vs. $M_{unhappy}$ = 6.39). All subsequent analyses focus on the primary anticipated emotional experience for the outcome in order to avoid floor effects for ratings of non-primary experiences (i.e., low ratings of unhappiness for those who received the grade they expected or a higher grade; low ratings of happiness for those who received a grade lower than they expected).

To determine the extent to which people differentiate specific features in their reports of forecast and experienced emotion, we examined the correlations among ratings of the intensity of emotion about an event, frequency of emotion about an event, and the effect of an event on mood. The correlations among the intensity of emotion, the frequency of emotion, and effect on mood were moderate at forecast ($r_{\text{int*freq}}$ = .54, *95% CI* [.48, .59]; $r_{\text{int*mood}}$ = .48, *95% CI* [.42, .54]; $r_{\text{freq*mood}}$ = .55, *95% CI* [.49, .60]), with an average overall correlation coefficient of .50.

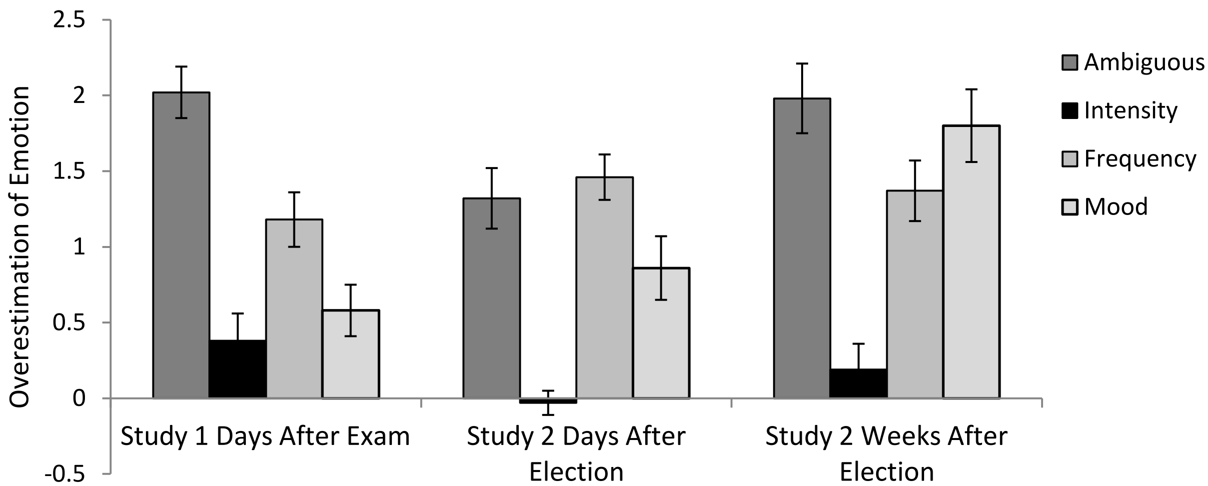
This means that the coefficient of determination (i.e., the proportion of variance in reports of one emotion feature that can be attributed to the other emotion feature) was an average of 25%. Correlations among features were also moderate for reports of emotional experience, with an average correlation of .54 ($r_{int*freq}$ = .54, *95% CI* [.48, .59]; $r_{int*mood}$ = .56, *95% CI* [.51, .61]; $r_{freq*mood}$ = .63, *95% CI* [.58, .67]), giving an average of 29% shared variance among reports of features of emotion. Given past work demonstrating that people have greater difficulty estimating emotional experiences at forecast than at experience (e.g., Wilson & Gilbert, 2005), it was possible that people might distinguish features of emotion less when forecasting than when reporting their emotional experience. A comparison of the average correlation for these features of emotion at forecast (.50) versus at experience (.54) revealed no significant difference in the size of the relationship among features of emotion at these two time points, $z = 1.70$, $p = .09$, *Cohen's q* = .06. Although not the focus of this investigation, correlations of "ambiguous emotion" with other emotion features were, on average, stronger at forecast ($r = .55$) than at experience ($r = .41$), $z = 5.59$, $p < .001$, *Cohen's q* = .18. In sum, the moderate correlations among reports of emotion support the proposition that people make distinctions among the intensity of emotion, frequency of emotion, and effect on mood, with approximately 71-75% of variance in reports of each feature of emotion not explained by the other two features.

Bias in Forecasts

To assess the degree of bias across features of emotional experience, we conducted an Analysis of Variance (ANOVA) on forecast minus experienced emotion with emotion feature (ambiguous, intensity, frequency, mood) as the repeated measures factor. The dependent variable was degree of overestimation (forecast – experienced emotion); therefore a positive mean indicates overestimation of future emotion (i.e., forecast emotion was greater than experienced emotion). This omnibus test is important, in part, because it accounts for the fact that measurements were repeated within the same participant. As Figure 2 shows, the degree of overestimation differed by emotion feature, $F(3, 1797) = 95.58$, $p < .001$, $\eta_p^2 = .14$. Follow up post-hoc contrasts revealed that participants were more biased in forecasting "ambiguous emotion" ($M = 2.02$, $SD = 2.18$, *95% CI* [1.85, 2.19]) compared to the intensity of their emotion, $t(608) = 15.40$, $p < .001$, $d = .64$, *95% CI* [.52, .75], the frequency of their emotion, $t(604) = 7.50$, $p < .001$, $d = .31$, *95% CI* [.20, .42], and the effect of the exam grade on their overall mood, $t(611) = 13.14$, $p < .001$, $d = .53$, *95% CI* [.42, .64], consistent with past research showing the question format for "ambiguous emotion" can inflate inaccuracy (Levine et al., 2012).

More critically, consistent with hypotheses, participants were less biased in forecasting the intensity of their emotion ($M = 0.36$, $SD = 2.27$, *95% CI* [0.18, 0.54]), compared to the frequency of emotion ($M = 1.16$, $SD = 2.27$, *95% CI* [0.98, 1.34]), $t(619) = 7.52$, $p < .001$, $d = .30$, *95% CI* [.19, .42]. They were also less biased in forecasting the intensity of emotion compared to the effect of the exam on overall mood ($M = 0.55$, $SD = 2.16$, *95% CI* [0.38, 0.72]), $t(629) = 1.97$, $p = .049$, $d = .08$, *95% CI* [-.03, .19]. Although not predicted, there was also less forecasting bias for effect on mood than emotion frequency, $t(622) = 6.03$, $p < .001$, $d = .25$, *95% CI* [.14, .36].

Figure 2. Overestimation in forecasting different features of emotion in Studies 1 and 2.



Note. Overestimation of emotion refers to forecast emotion minus experienced emotion (95% *CI* in bars).

One-sample *t*-tests were used to compare forecasting bias to zero (indicating no difference between forecast and experience, i.e., accuracy; see above for associated descriptive statistics). These comparisons revealed that, consistent with past work, participants overestimated "ambiguous emotion" (i.e., how they would feel in general after receiving their grade), $t(613) = 23.01$, $p < .001$. Participants also overestimated the intensity of their emotions about the grade, $t(631) = 3.96$, $p < .001$, although, as described above, participants were significantly less biased in estimating the intensity of emotion than any other emotion feature. Consistent with hypotheses, participants overestimated the frequency of their emotional responses to their grade, $t(624) = 12.75$, $p < .001$, and the effect of the exam grade on their overall mood, $t(640) = 6.44$, $p < .001$.

Predictors of Bias

Study 1 included measurements of the cognitive features that we theorized would contribute to forecasting biases for different features of emotion, as a preliminary examination of the relevance of these proposed mechanisms. Causal inferences, however, can only be made from the laboratory studies (Studies 3-5). Separate linear regressions were conducted to examine the relationship between potential predictors of bias and the degree of bias in forecasts of each feature of emotion (intensity, frequency, mood), as well as for forecasts of "ambiguous emotion." These analyses included three potential predictors of bias in forecasts: changes in frequency of thinking about the event from forecast to experience, changes in the appraised importance of the event from forecast to experience, and changes in surprisingness of the event from forecast to experience. As shown in Table 1, misestimation of the frequency of thinking about the exam grade predicted misestimation in forecasts of intensity, frequency, and mood. Changes in appraisals of surprise also predicted misestimation in forecasts of intensity, frequency, and mood. Changes in appraisals of importance did not predict bias for any feature of emotion (but see subsequent studies). The inferences from these analyses remain identical, although the standardized beta coefficients are on average smaller, if bias in the other features of emotion are included as covariates (e.g., bias in intensity of emotion and bias in frequency of emotion included as predictors of bias in effect on mood).

Table 1. Linear regressions predicting bias in forecasts of ambiguous emotion, emotion intensity, emotion frequency, and effect on mood in Study 1.

	B	B [*95% CI*]	*t*	*p*
Ambiguous emotion bias $R^2 = .05, F(3, 602) = 9.62, p < .001$				
Difference in thought frequency	.21	.20 [.12, .27]	5.11	<.001
Difference in importance	.04	.03 [-.04, .10]	0.87	.386
Difference in surprise	-.01	-.01 [-.08, .06]	-0.22	.826
Intensity bias $R^2 = .09, F(3, 621) = 20.73, p < .001$				
Difference in thought frequency	.16	.16 [.09, .24]	4.14	<.001
Difference in importance	.02	.02 [-.05, .08]	0.44	.659
Difference in surprise	.22	.20 [.13, .27]	5.60	<.001
Frequency bias $R^2 = .27, F(3, 614) = 75.72, p < .001$				
Difference in thought frequency	.47	.46 [.39, .53]	13.13	<.001
Difference in importance	.04	.04 [-.03, .10]	1.17	.243
Difference in surprise	.14	.13 [.06, .19]	3.87	<.001
Mood bias $R^2 = .18, F(3, 628) = 48.27, p < .001$				
Difference in thought frequency	.41	.39 [.32, .46]	11.18	<.001
Difference in importance	.00	.00 [-.06, .06]	0.43	.966
Difference in surprise	.07	.06 [-.00, .12]	1.84	.066

Note. Multicollinearity statistics were as follows in models predicting one cognitive feature with the other two cognitive features: difference in thought frequency (tolerance = .99, VIF = 1.01), difference in importance (tolerance = .95, VIF = 1.05), difference in surprise (tolerance = .99, VIF = 1.01).

Follow up comparisons showed that the strength of the associations between cognitive features and bias differed by feature of emotion. Compared to intensity bias, misestimation of the frequency of thinking about the exam grade was a stronger predictor of frequency bias, $z = 6.13$, $p < .001$, *Cohen's q* = .35, and mood bias, $z = 4.85$, $p < .001$, *Cohen's q* = .27. Changes in the surprisingness of the exam grade between forecast and experience predicted intensity bias more strongly than it predicted mood bias, $z = 2.71$, $p = .007$, *Cohen's q* = .15. However, causality is difficult to address in a correlational design, and changes in frequency of thought, perceived importance, and perceived suprisingness tended to be correlated ($r_{freq*surp}$ (636) = .22, p < .001, *95% CI* [.15, .29]; $r_{freq*imp}$ (638) = .11, p = .006, *95% CI* [.03, .19]; $r_{surp*imp}$ (637) = .08, p = .05, *95% CI* [.003, .16]). Thus the laboratory studies (Studies 3-5) that experimentally manipulated cognitive features provide the strongest causal evidence.

Study 2: 2016 U.S. Presidential Election

Study 1 demonstrated varying degrees of bias in forecasts of the intensity of emotion about an event, the frequency of emotion about an event, and the effect of an event on overall mood. In Study 2 we sought to replicate these findings using a large-scale national event similar to that assessed in previous investigations of affective forecasting – a presidential election (e.g., Levine et al., 2012; Wilson et al., 2003). Experienced emotion was assessed twice in this study, permitting an examination of changes in forecasting biases over time since an event. As described above, we hypothesized that bias would increase over time for forecasts of the effect of an event on mood, as people gradually adjust to and cope with an event after it occurs. Three weeks prior to the election, Study 2 participants provided forecasts for two time points: the week that the election outcome was announced and three weeks later, and they later reported their experience at these two time points.

Method

Participants. Data collection was planned for approximately 1,000 participants, based on estimates that this would provide sufficient power to detect a small effect size. (An a priori g*power computation for the difference between two dependent means gave a total sample size of 327 to have power of .95 to detect a d effect size of .20. An a priori g*power computation for a MANOVA with 2 repeated measures and 2 groups gave a total sample size of 860 to have power of .95 to detect an effect size f of .10.) We oversampled due to anticipated attrition over time points, for a total of 1,183 participants who completed the first survey. Participants were offered a payment (for MTurk) or course credit (for students) for each survey they completed.

Participants were excluded if they did not answer questions at subsequent time points (n = 146), indicated that the election of either candidate would be a negative event (n = 313) or that a victory by either candidate would be a positive event (n = 18). Excluding these groups makes the results of our study more comparable to past research on forecasting biases because the participants included had experienced an outcome that they clearly preferred (information for the group who viewed both major candidates as negative is available in Supplemental materials).

The final sample included 706 participants, with 346 participants recruited through Amazon Mechanical Turk (Mturk; Buhrmester, Kwang, & Gosling, 2011), 110 undergraduate students enrolled in a large public Texas university, and 250 undergraduate students enrolled in a large public California university. This recruitment provided a post hoc power of .9996 to detect a .20 *d* effect size. Different groups were sampled to capture participants that varied in their political preferences and other demographic characteristics. The sample included 67% women, 32% men, and .8% with other gender identities, with an average age of 29.45 years (SD = 12.71). Of the 86% of participants who reported voting, 70% voted for Clinton, 28% for Trump, and 3% for another candidate.

Procedures and materials. This study was part of a larger investigation, and only methods and procedures relevant to the present analyses are reported here. Participants completed online surveys at three time points: three weeks before the election, the week of the election, and three weeks after the election. Three weeks before the election, participants forecast how they would feel the week of the election and three weeks after the election. At those two later time points, participants reported their current feelings. Participants forecast how they would feel in the evening, and reported their emotional experience in the evening (for their time zone) to provide consistency between reports of forecast and experience.

Forecasts (Time 1). Participants were prompted to "Suppose it is an evening during the week of November 8th, days after the presidential election, and that Donald Trump won the election and will be the next president of the United States." After this prompt, participants were asked: "In general, how will you be feeling at that time?" They forecast how happy, angry, and scared they would feel in general on scales ranging from *not at all* (1) to *extremely* (9). As in Study 1, we limit inferences based on this "ambiguous emotion" question because participants often misunderstand the question (Levine et al., 2012).

Participants then forecast their feelings using questions that focused specifically on their response to the event, including the intensity of their emotion, the frequency of their emotion, and the effect of the event on mood. For intensity, they were asked: "How will you be feeling about Donald Trump being elected president? How intensely will you feel: Happy, Angry, Scared" on scales ranging from *not at all* (1) to *extremely* (9). They also forecast the frequency of their feelings: "How frequently that day will you feel this way about Donald Trump being elected president? Happy, Angry, Scared" on scales from *not at all - 0%* (1) to *the entire day - 100%* (11) in 10-point percentage increments. Participants forecast the effect of the event on their mood: "Overall, how much of the time that day will you spend in a mood that is…Happy, Angry, Scared" on the same percentage scale used for frequency. Using the same procedure, participants forecast how they would feel during an evening three weeks after the election.

As in Study 1, the primary question addressed by Study 2 was the degree of overestimation in forecasts of different features of emotion, but again we included a preliminary assessment of the relevance of three cognitive features theorized to contribute to forecasting biases. Participants forecast how frequently they would think about the election an evening the week of the election and an evening three weeks after the election, “Overall, how much of the time that day will you spend thinking about Donald Trump being elected president” using the same percentage scale as for frequency of feeling. Participants also reported their appraisals of the election outcome. They reported their appraisal of the importance of the election, “How important is the outcome of the 2016 presidential election to you?” on a scale ranging from *not at all important* (1) to *extremely important* (9). They also forecast, for an evening the week of the election and three weeks later, “How surprised will you be if Donald Trump is elected president?” on a scale from *not at all surprised* (1) to *extremely surprised* (9). Participants answered two dichotomous questions by reporting whether a Trump victory would be good or bad and whether a Clinton victory would be good or bad.

Experience the week of the election (Time 2). Participants who completed the Time 1 survey were invited to complete the second survey any time between 5 p.m. and midnight, from November 9th-12th, 2016. As a pair to the “ambiguous emotion” question asked at forecast, consistent with the typical affective forecasting procedure, participants first reported their current feelings with no reference to the election outcome, “In general, how are you currently feeling? Happy, Angry, Scared” on scales ranging from *not at all* (1) to *extremely* (9).

They were then reminded that, “Donald Trump won the election and will be the next president of the United States,” and asked to report their emotional experiences. All emotional experience questions about the intensity of emotion, frequency of emotion, and effect on overall mood used the same questions and scales as at forecast, with verbs changed from future to current tense. Participants reported their current appraisals of the importance of the election outcome, as well as their frequency of thinking about the election outcome, and their degree of surprise at the election outcome, using the same questions as at forecast. They also reported whether or not they had voted and, if they had voted, the candidate they voted for.

Time 3: Three weeks after the election. Participants who completed the Time 1 survey were invited to complete the third survey any time between 5 p.m. and midnight from November 28th through December 2nd, 2016. Participants reported their experienced feelings and current appraisals using the same procedure, questions, and scales as at Time 2.

Results

Preliminary analyses. A total of 184 participants (26%) viewed a Trump victory as a positive event and a total of 522 participants (74%) viewed a Trump victory as a negative event. All subsequent analyses focus on the primary emotional experience for type of event (i.e., happiness for those who perceived a Trump victory as a positive event; mean ratings of angry and scared for those who perceived a Trump victory as a negative event). Ratings of angry and scared were highly correlated at all time points and therefore combined for analyses.

To determine the extent to which participants differentiated specific features in their reports of forecast and experienced emotion, we examined the correlations among these features of emotion. The correlations among the intensity of emotion, the frequency of emotion, and effect on mood were strong at forecast for days after the election ($r_{\text{int*freq}} = .78$, *95% CI* [.75, .81]; $r_{\text{int*mood}} = .70$, *95% CI* [.66, .74]; $r_{\text{freq*mood}} = .83$, *95% CI* [.81, .85]), with an average overall correlation coefficient of $r = .78$. Similarly, correlations among reports were strong at forecast for weeks after the election ($r_{\text{int*freq}} = .86$, *95% CI* [.84, .88]; $r_{\text{int*mood}} = .55$, *95% CI* [.50, .60]; $r_{\text{freq*mood}} = .64$, *95% CI* [.60, .68]), with an average overall correlation coefficient of $r = .68$. Thus, an average of 61% of variance in reports of one feature of emotion could be explained by variance in the other two features for forecasts days after the election, and an average of 46% of the variance in reports of one feature of emotion could be explained by other features for forecasts three weeks after the election.

The week after the election, the average overall correlation coefficient among experienced intensity, frequency, and mood was r = .68 ($r_{\text{int*freq}}$ = .79, *95% CI* [.76, .82]; $r_{\text{int*mood}}$ = .62, *95% CI* [.57, .66]; $r_{\text{freq*mood}}$ = .69, *95% CI* [.65, .73]), giving an average of 46% shared variance among reports of features of emotion. Three weeks after the election, the average overall correlation coefficient among reports of features of emotion was r = .51 ($r_{\text{int*freq}}$ = .65, *95% CI* [.60, .69]; $r_{\text{int*mood}}$ = .49, *95% CI* [.43, .55]; $r_{\text{freq*mood}}$ = .50, *95% CI* [.44, .56]), giving an average of 26% shared variance among reports of features of emotion.

As noted in Study 1, previous theory suggests that people might distinguish among features of emotion more at experience than they do at forecast. To examine this possibility, a comparison was conducted on the average correlation among features at forecast versus experience; coherence was significantly weaker at experience the week after the election than at forecast, z = 4.05, p < .001, *Cohen's q* = .22. The relationship among features was also significantly weaker at experience three weeks after the election than at forecast, z = 4.89, p < .001, *Cohen's q* = .27. Further, the relationships among features of emotion were significantly weaker for experience three weeks after the event than the week of the election, z = 4.91, p < .001, *Cohen's q* = .27. Thus, differentiation of features of emotion appeared to increase as time elapsed from the occurrence of the event.

Correlations of "ambiguous emotion" with other emotion features were also stronger at forecast (r = .71) than at experience the week of the election (r = .65), z = 3.63, p < .001, *Cohen's q* = .11, which were stronger than correlations at experience three weeks later (r = .54),

$z = 5.45$, $p < .001$, *Cohen's q* = .17. Thus, people appear to differentiate features of emotional states less at forecast than they do at experience.

In sum, the moderate correlations among reports of emotion support the proposition that people make distinctions among features of emotion (intensity, frequency, mood), even at the time of forecasting. At forecast, 46%-61% of variance was shared among reports of features of emotion, decreasing to 46% for experience days after the election, and then to 26% for experience three weeks after the election. Thus approximately 39-74% of variance in reports of each feature of emotion was not explained by other features of emotion. The pattern of results further suggests that, at least for some events, people differentiate features of emotion to a greater extent after they experience an event than when they are forecasting how they will react to a future event.

Bias in forecasts. To compare the degree to which participants overestimated different features of their emotional experience over time, we conducted a mixed model ANOVA on forecast minus experienced emotion. The repeated measures were emotion feature (ambiguous, intensity, frequency, mood) and time point (week of election, three weeks later). This omnibus test is important, in part, because it accounts for the fact that measurements were repeated within the same participant. This analysis revealed a main effect of emotion feature, $F(3, 1893) = 154.57$, $p < .001$, $\eta_p^2 = .20$, a main effect of time point, $F(1, 1893) = 29.14$, $p < .001$, $\eta_p^2 = .04$, and an interaction between emotion feature and time point, $F(3, 1893) = 23.56$, $p < .001$, $\eta_p^2 = .04$.

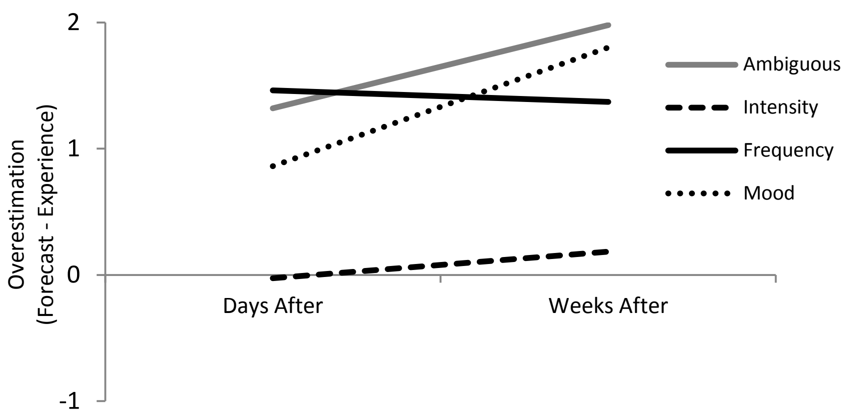
The main effect of emotion feature indicated that the degree of bias differed across features of emotion. In these analyses, a positive mean indicates overestimation of future emotion (i.e., forecast emotion was greater than experienced emotion). Follow up post-hoc contrasts revealed that, as in Study 1, participants showed greater bias in rating their "ambiguous emotion" (*M* = 1.62, *SD* = 2.33, *95% CI* [1.45, 1.79]), compared to the intensity of their emotion, *t*(705) = 20.77, *p* < .001, *d* = 0.72, *95% CI* [.61, .82], the frequency of their emotion, *t*(705) = 2.99, *p* = .003, *d* = 0.11, *95% CI* [.004, .21], and the effect of the election on overall mood), *t*(704) = 4.66, *p* < .001, *d* = 0.19, *95% CI* [.08, .29].

As shown in Figure 2, consistent with hypotheses, participants overestimated less in forecasting the intensity of emotion (*M* = .04, *SD* = 1.78, *95% CI* [-.09, .17]) compared to forecasting the frequency of emotion (*M* = 1.37, *SD* = 2.16, *95% CI* [1.21, 1.53]), *t*(705) = 20.44, *p* < .001, *d* = 0.64, *95% CI* [.77, .99]. Participants also overestimated less in forecasting the intensity of emotion compared to forecasting the effect on overall mood (*M* = 1.25, *SD* = 2.54, *95% CI* [1.06, 1.44]), *t*(704) = 14.76, *p* < .001, *d* = 0.71, *95% CI* [.60, .82]. The degree of overestimation did not significantly differ for forecasts of emotion frequency and mood, *t*(704) = 1.64, *p* = .10, *d* = 0.07, *95% CI* [-.04, .17].

One-sample *t*-tests were used to compare the degree of bias in forecasts to zero (which represents no difference between forecast and experienced emotion; i.e., accuracy). These comparisons revealed that participants overestimated their "ambiguous emotion" after the election, *t*(705) = 18.40, *p* < .001. Participants were accurate in forecasting the intensity of their future emotions about the election, *t*(705) = 0.59, *p* = .55. Participants overestimated the frequency of their emotional responses to the election, *t*(705) = 16.74, *p* < .001, and the effect of the election on their overall mood, *t*(704) = 13.02, *p* < .001.

The design of this investigation also permitted an examination of changes in forecasting bias over time since an event. As reported above, there was an interaction in the omnibus test between feature of emotion and time point (days after versus weeks after the election). Follow up contrasts were conducted to determine if the degree of overestimation for features of emotion (i.e., the difference between forecast and experienced emotion) differed the week of the election versus three weeks later. The pattern of results is shown in Figure 3. Although not the primary focus of this investigation and included only for reference to the larger forecasting literature, participants overestimated their "ambiguous emotion" to a greater degree three weeks after the election (M = 1.98, SD = 2.82, *95% CI* [1.75, 2.19]) than they did a few days after the election (M = 1.32, SD = 2.72, *95% CI* [1.12, 1.52]), $t(647) = 5.28$, $p < .001$, d = 0.20, *95% CI* [.10, .31]. The degree of overestimation in forecasts did not significantly change over time for intensity of emotion (M_{T2} = 0.04, SD = 1.88, *95% CI* [-.16, .12]; M_{T3} = 0.18, SD = 2.24, *95% CI* [.01, .35]), $t(651) = 1.62$, p = .106, d = .07, *95% CI* [-.04, .18], or for the frequency of emotion (M_{T2} = 1.44, SD = 2.51, *95% CI* [1.21, 1.59]; M_{T3} = 1.36, SD = 2.56, *95% CI* [1.16, 1.56]), $t(648) = 0.72$, p = .472, d = 0.03, *95% CI* [-.08, .14]. Consistent with hypotheses, participants overestimated the effect of the election on mood to a greater degree a few weeks after the election (M = 1.76, SD = 3.07, *95% CI* [1.52, 2.00]) than they did a few days after the election (M = 0.84, SD = 2.88, *95% CI* [.61, 1.04]), $t(649) = 7.11$, $p < .001$, d = 0.27, *95% CI* [.16, .38]. Thus, participants overestimated the effect of the election on their mood, and this bias increased when they were forecasting reactions more distal from the event.

Figure 3. Overestimation in forecasts of different features of emotion days and weeks after an event.



Predictors of bias. As with Study 1, measurements of the cognitive features that we theorized would contribute to forecasting biases for different features of emotion were examined, as a preliminary indicator of the relevance of these proposed mechanisms. Causal inferences, however, can only be made from the laboratory studies (Studies 3-5). Separate linear regressions were conducted to examine the relationship between cognitive features and the degree of bias in each feature of emotion. These analyses were conducted for intensity, frequency, and effect on mood, as well as “ambiguous emotion.” However, we limit inferences from the “ambiguous emotion” question, as noted previously. These analyses controlled for bias in other features of emotion as covariates (e.g., in the analysis examining bias in mood, intensity bias and frequency bias were included as covariates; results without covariates are reported in supplemental materials because inclusion of these covariates arose during the review process and was not specified a priori). As shown in Table 2, consistent with hypotheses, misestimation in forecasts of the frequency of thinking about the election predicted misestimation in forecasts of frequency of emotion about the election, and further predicted bias in forecasting the effect of the election on mood. Consistent with hypothesized relationships, changes in appraisals of importance predicted bias in forecasts of intensity of future emotion. Changes in appraisals of surprise did not relate to forecasting biases, except for intensity bias three weeks after the election outcome.

Table 2. Linear regressions predicting bias in forecasts of ambiguous emotion, emotion intensity, emotion frequency, and mood in Study 2.

	Week of the Election				**Three Weeks after Election**			
Emotion feature	β	B *[95% CI]*	*t*	*p*	β	B *[95% CI]*	*t*	*p*
Ambiguous emotion	R^2 = .37, F(6, 674) = 68.31, p < .001				R^2 = .43, F(6, 623) = 79.85, p < .001			
Dif. in thought freq.	.05	[-.03, .14]	1.36	.174	.05	[-.02, .13]	1.37	.172
Dif. in importance	.00	[-.08, .08]	-0.06	.950	.03	[-.04, .13]	1.09	.278
Dif. in surprise	.01	[-.06, .09]	0.44	.661	.06	[.01, .15]	2.11	.035
Intensity bias	R^2 = .78, F(5, 682) = 83.97, p < .001				R^2 = .09, F(3, 685) = 24.10, p < .001			
Dif. in thought freq.	-.06	[-.11, .01]	-1.77	.077	-.01	[-.08, .06]	-0.30	.763
Dif. in importance	.11	[.04, .15]	3.48	.001	.11	[.04, .18]	3.20	.001
Dif. in surprise	.04	[-.02, .09]	1.40	.163	.06	[.001, .13]	1.98	.048
Frequency bias	R^2 = .53, F(5, 677) = 154.86, p < .001				R^2 = .43, F(5, 631) = 98.31, p < .001			
Dif. in thought freq.	.17	[.12, .25]	5.86	<.001	.25	[.20, .33]	7.63	<.001
Dif. in importance	.01	[-.05, .07]	0.25	.806	.03	[-.03, .11]	1.07	.287
Dif. in surprise	.02	[-.04, .08]	0.58	.562	-.03	[-.10, .03]	-0.96	.336
Mood bias	R^2 = .42, F(5, 677) = 100.73, p < .001				R^2 = .32, F(5, 631) = 61.01, p < .001			
Dif. in thought freq.	.20	[.17, .33]	6.23	<.001	.17	[.13, .31]	4.71	<.001
Dif. in importance	.03	[-.04, .12]	1.07	.287	.03	[-.06, .13]	0.73	.466
Dif. in surprise	-.02	[-.09, .06]	-0.51	.611	.05	[-.02, .15]	1.47	.142

Note. Multicollinearity statistics were as follows in models predicting one cognitive feature with the other two features: Time 2: difference in thought frequency (tolerance = .99, VIF = 1.01), difference in importance (tolerance = .99, VIF = 1.00), difference in surprise (tolerance = .97, VIF = 1.03). Time 3: difference in thought frequency (tolerance = .99, VIF = 1.00), difference in importance (tolerance = 1.00, VIF = 1.00), difference in surprise (tolerance = .97, VIF = 1.03). Bias in ambiguous emotion bias controls for all features of emotion; biases in intensity, frequency, and mood each control for the other two features of emotion (but not ambiguous emotion, which is often misinterpreted as one of the other features).

We also compared the strength of associations among cognitive features (importance, frequency of thinking, surprisingness) and forecasting biases for each feature of emotion. The results showed that some cognitive features related to bias more strongly than others and that this relationship differed by feature of emotion. Specifically, misestimating the frequency of thinking about the election predicted bias in forecasting the frequency of emotion more strongly than bias in forecasting the intensity of emotion, Time 2: $z = 4.27$, $p < .001$, *Cohen's q* = .23, Time 3: $z = 4.72$, $p < .001$, *Cohen's q* = .27, and predicted bias in forecasting mood more than intensity, Time 2: $z = 4.84$, $p < .001$, *Cohen's q* = .26, Time 3: $z = 3.23$, $p < .001$, *Cohen's q* = .18. Changes in appraised importance did not predict intensity bias significantly more than frequency bias, Time 2: $z = 1.85$, $p = .06$, *Cohen's q* = .10, Time 3: $z = 1.43$, $p = .15$, *Cohen's q* = .08, or intensity bias more than mood bias, Time 2: $z = 1.48$, $p = .14$, *Cohen's q* = .08, Time 3: $z = 1.43$, $p = .15$, *Cohen's q* = .08. The relationship between changes in appraised surprise and bias did not differ by feature of emotion, all *ps* > .576. However, causality is difficult to address in a correlational design, and changes in frequency of thought, perceived importance, and perceived suprisingness tended to correlate with one another (Time 2: $r_{\text{freq*surp}}(700) = .04$, $p = .345$, *95% CI* [-.03, .11]; $r_{\text{freq*imp}}(693) = .18$, $p < .001$, *95% CI* [.11, .25]; $r_{\text{surp*imp}}(697) = .10$, $p = .010$, *95% CI* [.03, .17]; Time 3: $r_{\text{freq*surp}}(653) = .13$, $p = .744$, *95% CI* [.05, .20]; $r_{\text{freq*imp}}(647) = .18$, $p < .001$, *95% CI* [.11, .25]; $r_{\text{surp*imp}}(648) = .04$, $p = .38$, *95% CI* [-.04, .12]). Thus the laboratory studies (Studies 3-5) that experimentally manipulated cognitive features provide the strongest causal evidence.

Study 3: Laboratory Examination of Frequency of Thought

Studies 1 and 2 were conducted in the field, and as a result were well suited to address the degree of bias in forecasts of different features of emotion. The studies revealed that people on average show markedly less overestimation in predicting the intensity of their future emotion than in predicting the frequency of future emotion or the impact of events on future mood. These studies also included a preliminary examination of the relevance of several cognitive features that were theorized to contribute to forecasting biases. The findings supported the potential role of frequency of thinking about events, perceived importance of events, and perceived surprisingness of events in forecasting biases. There were indications that the relationships between these cognitive features and forecasting bias might vary by feature of emotion.

However, field studies do not permit the type of experimental control that is necessary for examining causal relationships. To address the role of cognitive features in predicting bias in forecasts, Studies 3, 4, and 5 were laboratory studies that experimentally manipulated cognitive features. To promote methodological rigor in these studies, each study: 1) included a paradigm that has previously shown large forecasting bias effects (Gilbert et al., 1998), 2) targeted a sample size for each condition that would permit detection of small/medium effects, 3) controlled laboratory conditions (extensive training of research assistants, standard and clear protocol for pre-study lab set up and participant interaction, floor marks for consistency of distance, reduction of noise/distractions), and 4) involved pilot testing of potential manipulations (for each study, four manipulations were developed of the target construct and the effects on manipulation checks were assessed before conducting the reported study).

Study 3 addressed the effect of changing frequency of thought on forecasting biases. We hypothesized that changes in the frequency of thinking about an event would be particularly relevant to bias in forecasting the frequency of emotion about that event. This mechanism – that bias in forecasts results when people think less about the event than anticipated, or “focalism” – has been a central explanation for forecasting bias (e.g., Gilbert et al., 1998; Wilson et al., 2000).

Past studies have demonstrated that people do not think about events as much as they believe they will when making forecasts, and that frequency of thinking predicts bias in forecasts (Wilson et al., 2000). In contrast to these previous explanations, based on our model that differentiates among features of emotion, we hypothesized that this mechanism would contribute more strongly to bias in forecasting the frequency of emotion than intensity or mood. This connection seems intuitive – the more you think about an event, the more you feel about an event – but this specific connection is not captured in the forecasting literature.

Method

Participants. Data collection was planned for approximately 150 participants, based on estimates that this would provide sufficient power to detect a small/medium effect size (an a priori g*power computation for the difference between two dependent means gave a total sample size of 147 to have power of .95 to detect a *d* effect size of .30). This estimated effect size was based on recent meta-analytic findings that the overall average effect size in the affective forecasting literature is of medium size (*g* of .55, with a *95% CI* from .42 to .68; Levine et al., 2012). Data collection stopped at the end of the semester, with a sample size of 138. Participants (n = 20) were excluded if they failed to follow the prompt, expressed strong suspicion, or did not have responses for forecast or experienced emotion. The majority of exclusions (n = 18) were necessitated by unexpected study disruptions, program failures, errors in procedure, and missing data due to program failure, participant choice, or experimenter error. If participants who had data available are included in analyses, inferences remain identical to those reported in text. The final sample included 118 participants who were recruited from introductory college courses for partial course credit. This recruitment provided a post hoc power of .90 to detect a .30 effect size. Although this was lower than ideal, this sample size provided a post hoc power of .99 to detect a

.42 effect size, which was the lower limit of the 95% confidence interval produced from a meta-analysis of the literature, as noted above. The sample included 77% women, with an average age of 18.43 years (*SD* = 0.73).

Procedures and materials. In order to elicit an emotional response within the lab setting, we utilized a scenario in which participants experienced rejection. This scenario had been previously used and generated a large effect size in previous studies of affective forecasting bias (Gilbert, Pinel, et al., 1998, Study 6; *g* = 1.06). Participants were told that local businesses were seeking students to report their opinions in a brief study for pay ($40), and they were working with the researchers to prescreen participants for that study. Participants were told that, if they were not selected, they would complete another task that takes more time than the paid study. As part of the prescreen process, participants were asked to respond to 10 questions into a microphone. They were told that three business students, located in an adjoining room, would evaluate their answers to determine whether they would be selected for the paid study. Potential participants would only be rejected if the business students unanimously concluded that the "applicant was unfit for the job." Participants were then told that they would complete a preliminary survey that, as part of the agreement with the local businesses, the researchers had been permitted to add prior to the prescreen task.

For the "preliminary survey," participants forecast how they would feel after learning whether they were selected for the paid study; forecasts were made for being chosen for the study and for not being chosen for the study. Participants were prompted to forecast the intensity of their feelings: "Suppose you are [not] chosen for the brief study with pay. Ten minutes after finding out, how will you feel about [not] being chosen for the study?" They forecast how happy and unhappy they would feel on scales ranging from *not at all* (1) to *extremely* (9). They responded to the “ambiguous emotion” question: “Ten minutes after finding out, how will you feel in general?” using the same scale. They also forecast the frequency of their happy and unhappy feelings, “Suppose it is ten minutes after you find out that you are not chosen for the brief study with pay. Tell us how you will answer the following question at that time: In the past 10 minutes, how frequently did you feel this way about not being chosen?” on scales from *not at all* (1) to *constantly* (9). Participants forecast the effect of their feelings on mood, “Ten minutes after finding out, how much of an impact will not being chosen have on your overall mood?” on a scale ranging from *no impact at all* (1) to *extremely impactful* (9).

We also assessed cognitive features. Participants forecast how frequently they would think about not being chosen, “Suppose it is ten minutes after you find out you were not chosen for the brief study with pay. Tell us how you will answer the following question at that time: In the past 10 minutes, how frequently did you think about not being chosen?” using the same scale as for frequency of feeling. Participants also reported the importance of the event, “How important is being chosen to you?” on a scale ranging from *not at all* (1) to *extremely* (9).

After completing the forecasting survey, participants were given the 10 questions for the prescreen task (e.g., “Why did you select your major?”) and reminded about the elements of the prescreen task (e.g., answer out loud, evaluations by business students). The research assistant then left participants alone in the room, and the participants answered the questions for three minutes into the microphone. After three minutes, the research assistant entered the room, stopped the participant, and indicated that the decision of the evaluators should be ready shortly. An associate then knocked, the research assistant left the room, and then returned with an envelope. The research assistant explained, “Here’s their decision,” and handed the envelope to the participant. Inside was a letter indicating that the participant was rejected for the additional study. The research assistant then asked participants to complete the alternative (unpaid) study.

Before completing the survey, participants were randomly assigned to one of three conditions intended to change the frequency with which they thought about the outcome of the prescreen task. Based on previous measurements and manipulations related to frequency of thought in the forecasting literature (e.g., Wilson et al., 2000), the manipulation focused on either distracting thought from the rejection or focusing thought on the rejection. As described in the study overview, several manipulations were pretested to identify conditions that changed frequency of thinking about the event and did not change other cognitive features. In the low frequency condition, participants explained their plans for the next day in detail. In the control condition, participants were instructed to relax for a few minutes (this is frequently used as a neutral condition in the emotion literature; Lench et al., 2011). In the high frequency condition, participants were instructed to explain in detail their thoughts about not being chosen for the study for pay.

The research assistant left participants to this task for exactly ten minutes, and then participants completed a second survey on which they reported their experienced reactions. The questions and scales were identical to those at forecast, with tense and formatting changed to capture current reactions.

Results

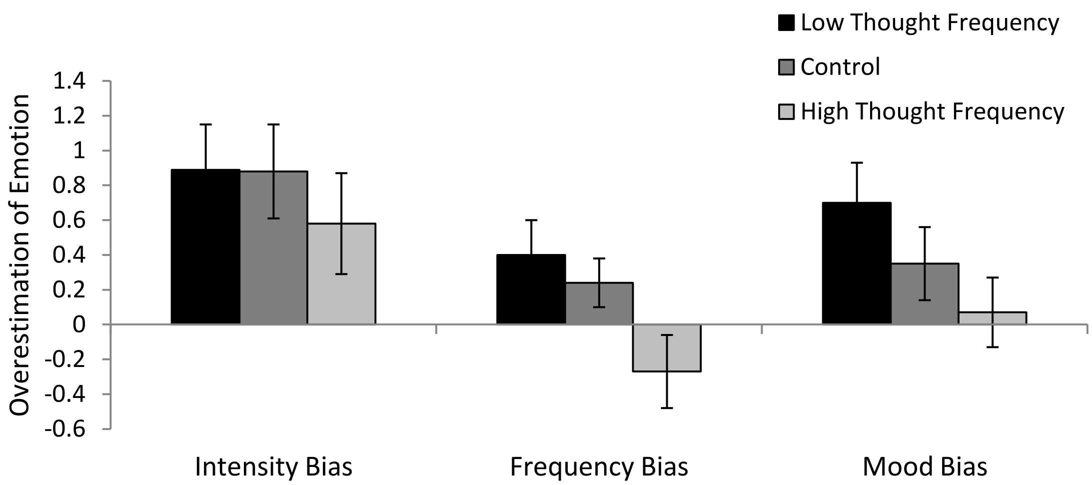
Preliminary analyses. To assess the effectiveness of the manipulation of thought frequency, we conducted ANOVAs with condition (low, control, high) as the between-subjects factor and the difference between forecast and experienced frequency of thought (forecast – experience) as the outcome. This analysis revealed that there was a difference in frequency of thought among conditions, $F(2,114) = 7.90$, $p = .001$, $\eta^2 = .122$, with participants in the high frequency condition reporting thinking about the event more often than they had forecast (M = -0.68, SD = 1.94, *95% CI* [-1.28, -.08]), and this difference was greater in the high frequency condition than in the control condition (M = 0.30, SD = 1.16, *95% CI* [-.06, .66]), $t(78) = 2.73$, p = .008, d = .61, *95% CI* [.16, 1.06], or the low frequency condition (M = 0.76, SD = 1.67, *95% CI* [.22, 1.30]), $t(75) = 3.46$, $p = .001$, $d = .79$, *95% CI* [.33, 1.26]. Overestimation of thought frequency did not significantly differ between the low frequency condition than the control condition, although the means were in the predicted direction, $t(75) = 1.40$, $p = .165$, $d = .32$, *95% CI* [-.13, .77], likely due to mind wandering during the relaxation period.

Similar analyses revealed that there were no differences in perceived importance among the conditions, $F(2, 115) = 1.72$, $p = .183$, $\eta^2 = .029$, and no differences in surprise among the conditions, $F(2, 114) = 1.96$, $p = .145$, $\eta^2 = .033$. Thus the manipulation effectively targeted frequency of thinking about the outcome, although the low frequency and control conditions did not significantly differ.

Effect of thought frequency on bias. To assess the relationship between thought frequency and bias in forecasts of features of emotion, we conducted separate ANOVAs for each emotion feature. In each analysis, thought frequency condition (low, control, high) was the between-subjects factor and overestimation in forecasting (forecast emotion – experienced emotion) was the outcome variable. Analyses that used repeated measures ANOVAs with forecast emotion and experienced emotion as the repeated measures yielded identical inferences to those reported below.

We hypothesized that the frequency of thinking about events was particularly relevant to bias in forecasts of the frequency of emotion about those events. Consistent with this hypotheses, thought frequency condition affected the degree of overestimation in forecasts of the frequency of feeling unhappy about the prescreen outcome, $F(2, 111) = 3.46$, $p = .035$, $\eta^2 = .059$. In these analyses, positive numbers represent overestimation of future emotion and negative numbers represent the underestimation of future emotion. As shown in Figure 4, consistent with hypotheses, post hoc comparisons revealed that overestimation in forecast frequency of emotion in the high thought frequency condition, where participants were prompted to focus on the event (*M* = -0.27, *SD* = 1.36, *95% CI* [-.69, .15]), was greater than in the low thought frequency condition (*M* = 0.40, *SD* = 1.19, *95% CI* [.01, .79]), $t(74) = 2.26$, $p = .027$, $d = .52$, *95% CI* [.06, .98], and was marginally greater than in the control condition (*M* = 0.24, *SD* = 0.88, *95% CI* [-.04, .52]), $t(77) = 1.94$, $p = .056$, $d = .44$, *95% CI* [-.01, .89]. Bias in forecasted frequency of feeling did not differ between the low frequency thought condition and the control condition, $t(71) = .67$, $p = .51$, $d = .15$, *95% CI* [-.30, .61], possibly because the manipulation did not effectively differentiate these two conditions. Thus, changing the frequency of thought caused changes in the extent to which participants overestimated the frequency of emotion.

Figure 4. Overestimation of emotion in forecasts after an experimental manipulation of thought frequency.



Note. Overestimation of emotion refers to forecast unhappiness minus experienced frequency of unhappiness. Standard error in bars.

Our hypotheses focused on comparisons among conditions, but we also examined the degree of bias evident in each condition compared to absolute accuracy. One-sample *t*-tests were used to compare the degree of bias in forecasts to zero (i.e., accuracy). Participants overestimated the frequency of their future unhappiness in the low frequency of thought condition, $t(34) = 1.98$, $p = .05$. They did not significantly overestimate the frequency of emotion in the control condition, $t(37) = 1.65$, $p = .107$, and did not significantly underestimate the frequency of emotion in the high frequency of thought condition, $t(40) = -1.26$, $p = .214$. Frequency of thought condition did not relate to bias in forecasts of "ambiguous emotion" after the prescreen outcome, $F(2, 105) = 1.43$, $p = .244$, $\eta^2 = .027$. Consistent with hypotheses, frequency of thought condition did not relate to bias in forecasts of the intensity of unhappiness about the prescreen outcome, $F(2, 113) = 0.42$, $p = .656$, $\eta^2 = .007$, and did not relate to bias in forecasts of the effect of the prescreen outcome on mood, $F(2, 115) = 2.14$, $p = .12$, $\eta^2 = .036$. Although patterns for other emotion features were generally in a similar direction as for frequency of emtion, changing frequency of thought did not significantly affect bias in forecasts of other emotion features. Therefore, an experimental manipulation of the frequency of thinking about an event was particularly relevant to bias in forecasts of the frequency of feeling about that event.

Study 4: Laboratory Examination of Importance of Event

Study 4 addressed the impact of the perceived importance of the event on forecasting biases. We hypothesized that the perceived importance of events was particularly relevant to bias in forecasting the intensity of emotion about that event. This hypothesis is based on findings from the emotion literature that the intensity of experienced emotion about an event is tightly coupled with the perceived importance of the event (Verduyn et al., 2009; 2012; 2013). Compared to the frequency of thought and the surprisingness of events, the importance of events for people's goals is fairly stable over time (McAdams & Olson, 2010), contributing to the accuracy in forecasts of intensity demonstrated in Studies 1 and 2. However, we anticipated that changing the importance of an event through experimental manipulation would cause changes in forecasting bias for the intensity of emotion.

Method

Participants. Study 4 used a paradigm similar to that employed in Study 3 and therefore estimated sample size was the same and data collection again ended at the end of the semester (with 154 participants). Participants (*n* = 19) were excluded if they failed to follow the prompt, expressed strong suspicion, or did not complete questions for forecast or experienced emotion. The majority of exclusions (*n* = 14) were necessitated by unexpected study disruptions or missing data. If participants who had data available are included in analyses, the inferences made remain identical to those reported in the text. The final sample included 135 participants who were recruited from introductory college courses for partial course credit. This recruitment provided a post hoc power of .93 to detect a .30 effect size; and a post hoc power of .998 to detect a .42 effect size, which was the lower limit of the 95% confidence interval produced from a meta-analysis of the literature, as noted above. Of the sample, 73% were women, and the average age of participants was 18.77 years (*SD* = 1.05).

Procedures and materials. As in Study 3, participants were told they would be prescreened for the opportunity to take part in a study for pay, and that a team of graduate students would evaluate their suitability based on answers to questions. They then completed the forecasting questions described in Study 3, found out that they were not selected to take part in the study, and completed the experience questions. Instead of the manipulation of thought frequency used in Study 3, participants were randomly assigned to one of three levels of a manipulation of the perceived importance of the prescreen outcome. The manipulation focused on changing the money associated with acceptance for the paid task, based on a large and established literature that uses monetary value to measure and manipulate importance of events (e.g., Lau, White, & Schnall, 2013; Menger, 1976). As described in the study overview, several manipulations were pretested to identify conditions that changed the importance of the event without changing other cognitive features. Importance was manipulated through the amount of incentive associated with the study for pay (from which participants were ultimately rejected). In the low importance condition, participants were told they could earn $5 in the study associated with the prescreen task. In the control/medium importance condition, participants were told they could earn $40 in the study. In the high importance condition, participants were told they could earn $100 in the study. Participants were provided this information during the initial task description, and reminded at the time that they received the rejection letter.

Results

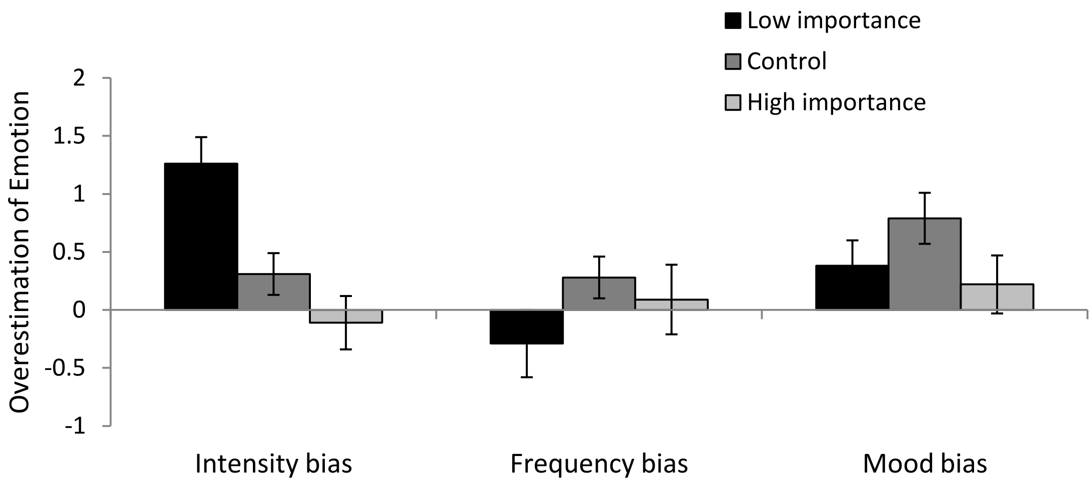
Preliminary analyses. To assess the effectiveness of the manipulation of importance, we conducted ANOVAs with condition (low, control, high) as the between-subjects factor and perceived importance of the rejection as the outcome. This analysis revealed a marginal difference in perceived importance among conditions, $F(2,132) = 2.98$, $p = .054$, $\eta^2 = .043$, with participants in the high importance condition perceiving the prescreen outcome as significantly more important ($M = 2.33$, $SD = 1.78$, *95% CI* [1.82, 2.84]), than participants in the low importance condition ($M = 1.64$, $SD = 1.01$, *95% CI* [1.33, 1.95]), $t(86) = 2.19$, $p = .031$, $d = .47$, *95% CI* [.05, .90]. Ratings of importance of the prescreen outcome in the control condition ($M = 1.83$, $SD = 1.17$, *95% CI* [1.50, 2.16]) fell between the high and low importance conditions, but did not significantly differ from the high importance condition, $t(91) = 1.60$, $p = .114$, $d = .33$, *95% CI* [-.08, .74], or the low importance condition, $t(87) = 0.80$, $p = .424$, $d = .17$, *95% CI* [-.24, .59].

Similar analyses revealed that there were no differences in frequency of thinking about the prescreen outcome among the conditions, $F(2, 132) = 2.14$, $p = .121$, $\eta^2 = .03$, and no differences in surprise among the conditions, $F(2, 131) = 1.61$, $p = .200$, $\eta^2 = .02$. Thus the manipulation resulted in changes in perceived importance of the prescreen outcome, particularly for the high versus low conditions, and did not significantly affect other cognitive features.

Effect of perceived importance on bias. To assess the relationship between perceived importance of the prescreen outcome and bias in forecasts of features of emotion, ANOVAs were conducted with importance condition (low, control, high) as the between-subjects factor and overestimation in forecast (forecast emotion – experienced emotion) as the outcome variable. Analyses that used repeated measures ANOVAs with forecast emotion and experienced emotion as the repeated measures yielded identical inferences to those reported below.

We hypothesized that perceived importance of events was particularly relevant to bias in forecasts of the intensity of emotion about those events. Importance condition affected the degree of overestimation in forecasts of the intensity of feeling unhappy about the prescreen outcome, $F(2, 125) = 10.48$, $p < .001$, $\eta^2 = .144$. Consistent with hypotheses, as shown in Figure 5, post hoc comparisons revealed that participants overestimated in forecasting emotional intensity less in the high importance condition ($M = -0.11$, $SD = 1.51$, *95% CI* [-.55, .33]) than in the low importance condition ($M = 1.26$, $SD = 1.41$, *95% CI* [.81, 1.71]), $t(81) = 4.26$, $p < .001$, $d = .94$, *95% CI* [.48, 1.39]. Although the contrast trended in the same direction, the degree of overestimation in forecasting emotional intensity was not significantly less in the high importance condition compared to the control condition ($M = 0.31$, $SD = 1.22$, *95% CI* [-.05, .67]), $t(88) = 1.46$, $p = .148$, $d = .31$, *95% CI* [-.11, .72]. Participants in the low importance overestimated emotional intensity more than did participants in the control condition, $t(81) = 3.30$, $p = .001$, $d = .73$, *95% CI* [.27, 1.17]. Thus, a manipulation of event importance caused changes in the degree of overestimation in forecasting the intensity of emotion.

Figure 5. Overestimation of emotion in forecasts after an experimental manipulation of perceived importance.



Note. Overestimation of emotion refers to forecast unhappiness minus experienced unhappiness. Standard error in bars.

Our hypotheses focused on comparisons among conditions, but we also examined the degree of bias evident in each condition compared to absolute accuracy. One sample *t*-tests examined whether the degree to which participants over- or underestimated the different features of emotion differed from zero, representing accuracy. Participants overestimated the intensity of their future unhappiness in the low importance condition, $t(37) = 5.53$, $p < .001$. Overestimation did not significantly differ from zero in the control condition, $t(44) = 1.71$, $p = .095$, or underestimation in the high importance condition, $t(44) = -0.49$, $p = .624$. Importance condition did not relate to bias in forecasts of “ambiguous emotion” after the prescreen outcome, $F(2, 114) = 1.14$, $p = .323$, $\eta^2 = .020$. Consistent with the proposition that importance of the event is particularly relevant to intensity bias, condition also did not relate to bias in forecasts of the frequency of feeling unhappy about the prescreen outcome, $F(2, 125) = 1.23$, $p = .296$, $\eta^2 = .019$, and did not relate to bias in forecasts of the effect of the prescreen outcome on mood, $F(2, 132) = 1.67$, $p = .192$, $\eta^2 = .025$. Therefore, an experimental manipulation of the perceived importance of an event was particularly relevant to bias in forecasts of the intensity of feeling about that event.

Study 5: Laboratory Examination of the Surprisingness of an Event

Study 5 addressed the impact of the perceived surprisingness of an event on forecasting biases. We did not initially hypothesize that surprise would relate to bias in forecasts, but included this cognitive feature in the field studies because of recent work demonstrating its relationship to the enduring effects of events on emotion (Verduyn et al., 2012; 2013). In Study 1, surprise predicted the magnitude of forecasting biases in several features of emotion. Based on this preliminary evidence that suggested surprise could be important to the processes under investigation, we included a laboratory investigation of the degree to which surprise caused bias in forecast of different features of emotion.

Method

Participants. Study 5 used a paradigm similar to that employed in Studies 3 and 4; therefore data collection estimates were the same, and data collection stopped at the end of the semester ($n = 139$). Participants ($n = 18$) were excluded if they failed to follow the prompt, expressed strong suspicion, or did not complete questions for forecast or experienced emotion. The majority of exclusions ($n = 14$) were necessitated by unexpected study disruptions, program failures, errors in procedure, and missing data due to program failure, participant choice, or experimenter error. If participants who had data available are included in analyses, the inferences made remain identical to those reported in text. The final sample included 121 participants who were recruited from introductory college courses for partial course credit. This recruitment provided a post hoc power of .91 to detect a .30 effect size. Although this is lower than ideal, this sample size provided a post hoc power of .996 to detect a .42 effect size, which was the lower limit of the 95% CI produced from a meta-analysis of the literature, as noted above. Of the sample, 80% were women, and the average age was 18.83 years (*SD* = 1.29).

Procedures and materials. As in Studies 3 and 4, participants were told they would prescreen for the opportunity to take part in a study for pay, and that a team of business students would evaluate their suitability based on answers to questions. They then completed the forecasting questions described in Study 3, found out that they were not selected to take part in the study, and completed experience questions. Participants were randomly assigned to one of three levels of a manipulation of the perceived surprisingness of the prescreen outcome. Surprise was manipulated through the likelihood of acceptance versus rejection for the paid task (all participants were ultimately rejected). Based on large literatures in psychology, public health, and economics that convey likelihood information in order to establish participants' expectation for outcomes, the manipulation verbally conveyed the likelihood of being accepted for the paid study during the initial task description. As described in the study overview, several manipulations were pretested to identify conditions that changed surprise at the outcome without influencing other cognitive features. In the low surprise condition, participants were told that almost nobody was accepted for the paid study and that they would probably be completing the alternative task. In the control condition, participants were not told any information about the likelihood of being chosen. In the high surprise condition, participants were told that almost everyone was accepted for the paid study and that they would probably be completing that task.

Results

Preliminary analyses. To assess the effectiveness of the manipulation of surprise, we conducted ANOVAs with condition (low, control, high) as the between-subjects factor and surprise after the event as the outcome. This analysis revealed a difference in surprise among conditions, $F(2, 116) = 35.90$, $p < .001$, $\eta^2 = .382$, with participants in the high surprise condition perceiving the prescreen outcome as more surprising ($M = 4.79$, $SD = 2.06$, *95% CI* [4.10, 5.48]), than participants in the low surprise condition ($M = 1.57$, $SD = 0.97$, *95% CI* [1.28, 1.86]), $t(74) = 9.01$, $p < .001$, $d = 2.07$, *95% CI* [1.51, 2.63]. Participants in the high surprise condition also perceived the rejection as more surprising than those in the control condition ($M = 3.21$, $SD = 1.82$, *95% CI* [2.67, 3.75]), $t(75) = 3.58$, $p = .001$, $d = .82$ *95% CI* [.35, 1.29]. Participants in the low surprise condition were less surprised than participants in the control condition, $t(83) = 5.16$, $p < .001$, $d = 1.12$, *95% CI* [.66, 1.58].

Similar analyses revealed that there were no differences in frequency of thinking about the prescreen outcome among the conditions, $F(2, 118) = 2.20$, $p = .120$, $\eta^2 = .04$, and no differences in importance among the conditions, $F(2, 117) = 1.32$, $p = .270$, $\eta^2 = .02$. Thus the manipulation resulted in changes in the surprisingness of the prescreen outcome and did not significantly affect other cognitive features.

Effect of surprisingness on bias. To assess the relationship between surprise of the prescreen outcome and overestimation in forecasts of features of emotion, ANOVAs were conducted with surprise condition (low, control, high) as the between-subjects factor and bias in forecast (forecast emotion – experienced emotion) as the outcome variable. Analyses that used repeated measures ANOVAs with forecast emotion and experienced emotion as the repeated measures yielded identical inferences as those reported below.

Surprise condition did not relate to bias in participants' forecasts of "ambiguous emotion" after the prescreen outcome, $F(2, 99) = 0.28$, $p = .754$, $\eta^2 = .006$, bias in forecasts of the intensity of unhappiness, $F(2, 114) = 1.18$, $p = .312$, $\eta^2 = .020$, bias in forecasts of the frequency of unhappiness after the outcome, $F(2, 115) = 0.22$, $p = .810$, $\eta^2 = .004$, and did not relate to bias in forecasts of the effect of the prescreen outcome on mood, $F(2, 117) = 0.17$, $p = .846$, $\eta^2 = .003$. Therefore, although the experimental manipulation of the surprisingness of the outcome was effective, surprise did not influence bias in forecasting any feature of emotion.

We conducted exploratory analyses across all studies to identify patterns that might explain why surprise mattered in Study 1 but did not influence forecasting bias in a laboratory manipulation. There was no evidence that these differences resulted from variation across studies in the valence of event or in correlations among cognitive features. The pattern of results did suggest that the relationship between surprise and emotional reactions differed across field versus laboratory studies. Ratings of emotion after the event were higher in the field studies ($M_{overall} = 5.93$) than in the laboratory studies ($M_{overall} = 2.54$). The correlations between surprise and ratings of emotion were relatively low and variable in the field studies (average $r = .10$), and higher and less variable in the lab studies (average $r = .43$). This pattern suggests that the relationship between surprise and emotional responses could differ for events that elicit strong versus weaker emotional responses. Study 5 experimentally manipulated surprise and demonstrated that surprise, in isolation, did not influence forecasting biases. Future studies should directly examine the features of events that vary the relationship between surprise and emotional responses.

Discussion

People's inability to forecast their future emotion is a primary barrier to making good decisions (e.g., Loewenstein, 2007). Most of the choices that people make, from inconsequential decisions about what food to order, to momentous decisions about who to marry, are based in part on how happy people think the outcomes of those decisions will make them. If people cannot accurately foresee how happy events will make them, they are virtually guaranteed to make bad decisions. Findings across several decades suggested that people were doomed to bad decisions – people could not foresee their future emotions, they consistently thought that future events would be better or worse than they actually turned out to be, and these biases were difficult or impossible to correct (Gilbert, 2006). Recent studies have shown that the landscape of forecasting biases is more complicated than these initial expositions. Studies have revealed mixed findings regarding people's ability to accurately forecast their emotions (e.g., Dore et al., 2016; Buechel et al., 2017), and a series of studies demonstrated that the most commonly used procedure to assess forecasting accuracy artificially inflates the magnitude of bias because participants misunderstand requests to forecast how they will feel in general (Levine et al., 2012).

This investigation provides a more nuanced perspective on people's strengths and weaknesses in forecasting how future events will make them feel. We proposed a theoretical framework that draws on the mechanisms that have been identified as contributors to forecasting bias as well as the broader emotion literature, to predict when and why people will show bias in forecasting emotion. The key component of this model is the differentiation among features of emotion, including the intensity of emotions in response to an event, the frequency of emotions about an event, and the effect of an event on mood. These features were selected because of their prominence in studies on forecasting and emotion theory, but there may be other features worth differentiating in future studies. This differentiated model of affective forecasting can account for the mixed findings in the forecasting literature, and yields insights into the psychological mechanisms that contribute to biases in forecasting emotion.

Differences in Bias

The first question addressed by this investigation was whether people are more accurate in forecasting some features of emotion than others. Based on past demonstrations of the cognitive features that contribute to different aspects of emotional experience, our model suggests that people should be better able to forecast the intensity of their future emotion than the frequency of their emotion or the effect of an event on mood. To test this prediction, forecast and experienced emotion was assessed for two real-world events that typically elicit strong emotional responses – the outcome of a college midterm exam and a U.S. Presidential election. Consistent with hypotheses, participants showed the least overestimation in forecasting the intensity of their future emotion about the events, and were consistently more biased in forecasting the frequency of emotion and the effect of the events on their mood. Further, examining bias over time for responses to the election revealed that bias was stable over time for intensity forecasts and frequency forecasts. In contrast, the tendency to overestimate the effect of the election on mood increased from a few days after the election to three weeks after the election. This finding is consistent with the proposition that forecasting bias for the effect of events on mood is particularly related to people's gradual adaptation to events after they occur. Overall, these findings demonstrate that the magnitude of bias depends on the specific aspect of emotional experience people are forecasting.

The theoretical zeitgeist for forecasting biases has been that people consistently exaggerate their future emotional reactions to events and several mechanisms were identified that contribute to this exaggeration. More recently, it has become clear that people sometimes overestimate, sometimes underestimate, and sometimes are accurate in forecasting their future emotions (e.g., Kaplan et al., 2016; Andrade & Van Boven, 2010; Buechel et al., 2018). As a result, there has been a growing lack of coherence in the field. A recent investigation demonstrated that the type of event being forecast can affect the size and direction of forecasting biases – events that elicit strong responses (e.g., important events, long events) tend to result in overestimation and events that elicit smaller reactions (e.g., less important events, short events) tend to result in underestimation (Buechel, Zhang, & Morewedge, 2017). Our findings indicate that, even for the same event, the feature of emotion that people are forecasting matters. People can more accurately foresee how intensely they will respond to an event than they can predict how frequently they will feel that way or how much the event will affect their mood. Further, our findings reveal the mechanisms that contribute to biases in forecasting specific features of emotion.

Mechanisms Contributing to Bias

Our model predicted that different mechanisms underlie bias in forecasting specific features of emotion. These proposed mechanisms were identified based on previous research showing the factors that determine the intensity of emotion, frequency of emotion, and the effect of events on mood (e.g., Gilbert et al., 1998; Levine et al., 2012; Verduyn et al., 2009, 2012). To test these mechanisms directly, we conducted a series of laboratory studies that experimentally manipulated cognitive features and assessed the consequences of these manipulations for forecasting biases. The results revealed that different mechanisms were particularly relevant to forecasting bias in different features of emotion. Specifically, manipulating the perceived importance of a negative event (being rejected from a study that offered payment) changed the degree of bias participants showed in forecasting how intensely unhappy they would be after the event. Participants in the low importance condition overestimated more in forecasting the intensity of unhappiness they would feel compared to participants in the control condition or high importance condition. This finding is consistent with past work showing that the intensity of people's emotional response to an event is strongly related to how important they perceive that event to be (Verduyn et al., 2009, 2012, 2013). Notably, manipulating importance did not affect bias in forecasting the frequency of unhappiness or the effect of the rejection on participants' mood. The importance ascribed to the event was particularly relevant to forecasts of the intensity of future emotion.

A separate laboratory study experimentally manipulated the frequency of thinking about an event. Thought content shifts rapidly over time, even after important events (Tully & Meyvis, 2017), and we hypothesized that people's inability to predict future thought content would underlie bias in forecasts of how frequently they would feel emotions about the event. This explanation – that people fail to anticipate future thought content or "focalism" – has been identified as a primary contributor to affective forecasting bias in the past (Wilson et al., 2000). Consistent with this view, experimental manipulation of the frequency of thought in past research did influence forecasting bias. In contrast to this past work, however, our model specifies that changes in the frequency of thought should promote bias particularly in forecasting the frequency of emotion. The results of the study revealed that changing the frequency of thinking about a negative event was particularly relevant to the magnitude of bias in forecasts of how frequently people would feel unhappy about that event in the future. Participants who were experimentally manipulated to think about a rejection less frequently overestimated more in forecasting how frequently they would feel unhappy relative to participants in the other conditions. The manipulation of thought frequency did not significantly change the magnitude of bias in forecasting the intensity of emotion or the effect of the rejection on mood. Thus, the present findings suggest a more nuanced view of this mechanism than has previously been theorized. The inability to predict thought content is particularly relevant to bias in forecasting the frequency of future feeling.

The results from these studies were framed in terms of overestimation, that is, the extent to which people forecast stronger emotion than they later experienced. However, it is important to note that changes over time in the appraised importance of events, or in the frequency of thinking about events, can promote underestimation of emotion. For example, people often appear to assume that future events will be more important than they later turn out to be (e.g., Gilbert & Wilson, 1995). At other times, however, people fail to appreciate the importance of future events, and underestimate the intensity of their emotional response as a result (e.g., Lench et al., 2011; Loewenstein, 2000). In such cases, further increasing the perceived importance of the experienced event should increase bias by leading people to underestimate the intensity of their emotional response even more. Therefore the reported findings should be used to make predictions about the degree of overestimation, which could result in greater or lesser accuracy depending on whether the direction of initial bias is toward over- or underestimation.

Consequences of Differential Bias

Given that people are far better at forecasting some features of their emotional experience than others, future investigations should assess how relying on forecasts of different features of emotion influences the quality of decision making. Theory and research provide reasons to expect that accurate forecasts of emotional intensity, frequency, and effects on mood are critical for effective decision making. Forecasts of how intensely one will respond to a future event provides an index of how good or bad that experience will be, whether one has the resources to cope with it, and the amount of effort it is worth expending to achieve or avoid it (Fredrickson, 2000). Indeed, for many decisions, the intensity of emotion people expect to feel, both while an experience is occurring and when they are thinking about it later, is the primary determinant of choice (e.g., musical performances, vacations, dentist visits, public speaking; Buehler & McFarland, 2001). Based on this evidence, the finding in the present investigation that people tend toward accuracy when forecasting the intensity of their future emotions bodes well for individual decision making.

Past theory and evidence also suggest that accurately forecasting the frequency of emotion and the effect of events on mood are important for decision making, and make a contribution separate from that of the intensity of emotion (e.g., Buehler & McFarland, 2001; Fredrickson, 2000; Wilson & Gilbert, 2005). These forecasts appear to provide an index of how much an event will dominate daily life compared to other events, and therefore inform how much people should prioritize achieving or avoiding that event over other goals. As Wilson and Gilbert (2005) have argued, overestimating an event’s consequences, by believing it will be salient and impactful in the future, can lead to poor investment of time and resources. As a result, people could expend effort and resources pursuing outcomes (e.g., new cars, retribution against a rival) that do not turn out to elicit the long-lasting happiness they anticipated, or working to avoid negative outcomes from which they will actually quickly recover.

Our finding of differential bias across features of emotion, suggests that reliance on various features of emotion to make decisions could have important consequences for the quality of those decisions. Specifically, people might benefit from the accuracy of forecasts of the intensity of emotion in identifying what choices they should make in relation to an event, such as pursuing an academic career based on the anticipation of intense happiness upon beginning a professorship. But overestimating the frequency of emotion and the effect of the event on mood could have detrimental consequences for decision making in the context of other events. For example, the pursuit of an academic career could be prioritized over social relationships and personal health because people overestimate the impact of attaining a professorship on daily happiness. In other words, the pattern of forecasting strengths and weaknesses observed suggest that people could accurately determine the value of an event in the moment while simultaneously overestimating the value of that event in relation to other events.

Forecasts of Individual Differences in Emotion

The present investigation focused on bias in affective forecasting using the definition most prevalent in the forecasting literature – assessing the degree to which people forecast more emotion than they later experienced. However, there are multiple ways to operationalize bias and accuracy in forecasting, and those different approaches can yield theoretically rich and informative results. Another way to conceptualize affective forecasting accuracy is to examine the correlation between forecast emotion and experienced emotion (Mathieu & Gosling, 2012). This approach reveals the degree to which people are relatively accurate – the extent to which people who predict they will respond more strongly later experience stronger emotion than people who predict they will respond less strongly. The mechanisms that we theorize as contributing to absolute bias (i.e., the difference between forecast and experienced emotion) are not necessarily relevant to relative accuracy.

As an example, the theoretical basis for the prediction that forecasts of intensity would tend toward accuracy was based on several established findings - including that importance of the event is relevant to experienced intensity of emotion and that perceived importance tends to be stable over time. It does not necessarily follow that any individual would anticipate that they are especially likely to view an event as more or less important than others, as this is a question about insight into the self. In fact, most people tend to exaggerate the importance of events for themselves (e.g., Blanton, Axsom, McClive, & Price, 2001; Gilovich, Medvec, & Savitsky, 2000). An interesting question for future research is the degree to which people can forecast their relative intensity of emotion, frequency of emotion, and effects on mood of events, and the mechanisms that contribute to bias in relative accuracy.

Concluding Statement

The picture of human decision making that has been conveyed in the affective forecasting literature is dire – people are inaccurate decision makers with little hope of attaining the happiness they seek despite their well-intentioned, but misguided, efforts to foresee the emotional consequences of their choices. Our findings suggest a more nuanced view of human decision making and of forecasting strengths and weaknesses. People’s ability to anticipate the intensity of emotion future events will evoke suggests that they are not doomed to make poor decisions, though they are vulnerable to several biases when trying to predict other features of their reactions to future events. Future studies will reveal whether there are strategies that can help people overcome these biases and make better decisions.

References

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

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, *36*, 917–927. https://doi.org/10.1037/0022-3514.36.8.917

Buechel, E. C., Zhang, J., & Morewedge, C. K. (2017). Impact bias or underestimation? Outcome specifications predict the direction of affective forecasting errors. *Journal of Experimental Psychology: General*, *146*, 746–761. https://doi.org/10.1037/xge0000306

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

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk. *Perspectives on Psychological Science*, *6*, 3–5. https://doi.org/10.1177/1745691610393980

Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., Miller, G. A., … Heller, W. (2003). Paying attention to emotion: An fMRI investigation of cognitive and emotional stroop tasks. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 81–96. https://doi.org/10.3758/CABN.3.2.81

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. https://doi.org/10.1006/jesp.2001.1481

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*, 276–302. https://doi.org/10.1037/0033-2909.125.2.276

Doré, B. P., Meksin, R., Mather, M., Hirst, W., & Ochsner, K. N. (2016). Highly accurate prediction of emotions surrounding the attacks of September 11, 2001 over 1-, 2-, and 7-year prediction intervals. *Journal of Experimental Psychology: General*, *145*, 788–795. https://doi.org/10.1037/xge0000168

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. https://doi.org/10.1037/0022-3514.92.6.990

Ebert, J. E. J., & Meyvis, T. (2014). Reading fictional stories and winning delayed prizes: The surprising emotional impact of distant events. *Journal of Consumer Research*, *41*, 794–809. https://doi.org/10.1086/677563

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion*, *14*, 577–606. https://doi.org/10.1080/026999300402808

Gilbert, D. T., Lieberman, M. D., Morewedge, C. K., & Wilson, T. D. (2004). The peculiar longevity of things not so bad. *Psychological Science*, *15*, 14–19. https://doi.org/10.1111/j.0963-7214.2004.01501003.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. https://doi.org/10.1037/0022-3514.75.3.617

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (2002). Durability bias in affective forecasting. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 292–312). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.018

Headey, B., & Wearing, A. (1989). Personality, life events, and subjective well-being: Toward a dynamic equilibrium model. *Journal of Personality and Social Psychology*, *57*, 731–739. https://doi.org/10.1037/0022-3514.57.4.731

Kaplan, R. L., Levine, L. J., Lench, H. C., & Safer, M. A. (2016). Forgetting feelings: Opposite biases in reports of the intensity of past emotion and mood. *Emotion*, *16*, 309–319. https://doi.org/10.1037/emo0000127

Kuppens, P., & Verduyn, P. (2017). Emotion dynamics. *Current Opinion in Psychology, 17,* 22-26.

Lench, H. C., & Bench, S. W. (2015). Strength of affective reaction as a signal to think carefully. *Cognition and Emotion*, *29*, 220–235. https://doi.org/10.1080/02699931.2014.904223

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. https://doi.org/10.1037/a0022792

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

Levine, L. J., Lench, H. C., Kaplan, R. L., & Safer, M. A. (2012). Accuracy and artifact: Reexamining the intensity bias in affective forecasting. *Journal of Personality and Social Psychology*, *103,* 584–605. https://doi.org/10.1037/a0029544

Levine, L. J., Lench, H. C., Kaplan, R. L., & Safer, M. A. (2013). Like Schrödinger's cat, the impact bias is both dead and alive: Reply to Wilson and Gilbert (2013). *Journal of Personality and Social Psychology*, *105,* 749–756. https://doi.org/10.1037/a0034340

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

Lucas, R. E., & Donnellan, M. B. (2007). How stable is happiness? Using the STARTS model to estimate the stability of life satisfaction. *Journal of Research in Personality*, *41*, 1091–1098. https://doi.org/10.1016/j.jrp.2006.11.005

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. https://doi.org/10.1177/0956797611427044

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

Morewedge, C. K., & Buechel, E. C. (2013). Motivated underpinnings of the impact bias in affective forecasts. *Emotion*, *13*, 1023–1029. https://doi.org/10.1037/a0033797

Suh, E., Diener, E., & Fujita, F. (1996). Events and subjective well-being: Only recent events matter. *Journal of Personality and Social Psychology*, *70*, 1091–1102. https://doi.org/10.1037/0022-3514.70.5.1091

Tully, S., & Meyvis, T. (2017). Forgetting to remember our experiences: People overestimate how much they will retrospect about personal event. *Journal of Personality and Social Psychology*, *113,* 878–891. https://doi.org/10.1037/pspa0000094

Verduyn, P., Delvaux, E., Van Coillie, H., Tuerlinckx, F., & Van Mechelen, I. (2009). Predicting the duration of emotional experience: Two experience sampling studies. *Emotion*, *9*, 83–91. https://doi.org/10.1037/a0014610

Verduyn, P., Van Mechelen, I., & Frederix, E. (2012). Determinants of the shape of emotion intensity profiles. *Cognition and Emotion*, *26*, 1486–1495. https://doi.org/10.1080/02699931.2012.662152

Verduyn, P., Van Mechelen, I., Tuerlinckx, F., & Scherer, K. (2013). The relation between appraised mismatch and the duration of negative emotions: Evidence for universality. *European Journal of Personality*, *27*, 481–494. https://doi.org/10.1002/per.1897

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting knowing what to want. *Current Directions in Psychological Science*, *14*, 131–134. https://doi.org/10.1111/j.0963-7214.2005.00355.x

Wilson, T. D., & Gilbert, D. T. (2008). Explaining way: A model of affective adaptation. *Perspectives on Psychological Science*, *3*, 370–386. https://doi.org/10.1111/j.1745-6924.2008.00085.x

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

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. https://doi.org/10.1037/0022-3514.78.5.821