Discrete and Dimensional Approaches to Affective Forecasting Errors

A dissertation submitted by

Prsni Patel

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Psychology

Tufts University

May 2025

Adviser: Heather L. Urry

Abstract

The process of making predictions about how one will feel in the future is known as affective forecasting. Most research on affective forecasting has focused on the idea that people make affective forecasting errors and predict more or less intense emotions than they actually experience. However, other evidence suggests that people can sometimes make accurate affective forecasts. In this research I examined one potential reason for this inconsistency – divergent measurement approaches – namely, discrete and dimensional. Specifically, through two studies, my goal was to examine whether affective forecasting errors differed based on whether they were measured using a discrete versus dimensional approach. In Study 1 ($N = 126$), I used data from an existing, within-subjects study. Participants viewed 20 descriptions of pictures that they would view in one week and made affective forecasts using both discrete and dimensional measures. One week later, participants viewed the pictures and rated how they actually felt using the same discrete and dimensional measures. Results suggested that affective forecasting errors did not differ based on measurement approach. I then conducted two studies in which I normed picture stimuli (Study 2a; $N = 50$) and picture descriptions (Study 2b; $N = 48$) to use in Study 3 ($N = 198$). Unlike in Study 1, Study 3 used stimuli normed to elicit discrete instead of dimensional affective states and a between-subjects design. Participants were randomly assigned to a discrete or dimensional condition. They saw 12 descriptions of pictures and made affective forecasts using discrete or dimensional measures. In the same session, they viewed the pictures and rated how they actually felt. Consistent with Study 1 and contrary to the hypothesis, affective forecasting errors did not differ based on measurement approach. Future research should examine whether these null results replicate, or whether there is a difference in forecasting errors based on measurement approaches in other contexts. Overall, these studies examined a novel source of variation in affective forecasting research and found that picture-based affective forecasting errors do not differ based on measurement approaches.

Keywords: affective forecasting errors, measurement approach, discrete emotions, dimensional affective states

Acknowledgements

I would not have been able to complete this dissertation without the help and support of numerous people.

To my dissertation committee - Dr. Elizabeth Race, Dr. Lisa Shin, and Dr. Ella Moeck. Each of you have had an impact throughout my time at Tufts. Liz, thank you for collaborating with Heather and I on our affective forecasting review paper and serving on my master's thesis committee (you're probably bored of hearing me talk about affective forecasting by now!). Lisa, thank you also for your feedback on my first year project and affective forecasting conceptual review paper - your input has always strengthened my work. Ella, thank you for being the only other person at SAS 2022 who studied affective forecasting and for having as much enthusiasm about this area as I did. I'm grateful to have had the opportunity to learn from the three of you over the last few years, and I'm excited to do so one last time.

To Kristen Petagna and Dr. Jolie Wormwood - thank you both for making the data and materials from your study openly available. Kristen, thanks also for graciously sharing your raw data with me, for answering all my questions, and enabling us to answer our research questions from an interesting, new angle.

To the scholars who took a chance on me, ignited my love for affective forecasting, and gave me my start in this field - Professor Allison Troy, Professor Michael Penn, Professor Christina Abbott, and Professor Megan Knowles - thank you for your support during my time at F&M.

To the friends I've made through the Psychology Department – Dr. Katie Ossenfort, and soon-to-be Drs. Julia Wefferling, Sanjana Kadirvel, McKinzey Torrance, and Monique Cathern –

thank you for commiserating when necessary, rotting on the couch with Taco Bell when necessary, and for making the last five years fly by as painlessly as possible.

To my boyfriend, Shree – these last three years have been replete with life's punches. Thank you for helping me not only roll with the punches but stand back up stronger (and ensuring I do so with a full stomach). Your support has meant the world to me.

To my grandparents – big nanu, small nanu, dadi, and dada - some of whom I lost during graduate school, thank you for your blessings in completing this Ph.D. I hope I made you proud.

To my family – ma, dad, beanie, gju. Beanu, thank you for listening to me blabber on about affective forecasting since my F&M days, for begrudgingly but nonetheless listening to each of my presentations on Zoom, and helping me navigate the inordinate amount of stress I had when scheduling this defense. I feel lucky to have you in my corner. Ma and dad, thank you for your unconditional support in every one of my undertakings, including this Ph.D. Regardless of whether you understood what I was talking about (which you always tried to!), you consistently showed up for me, despite being 7,681 miles away. Thanks for calling me every night at 9:30pm on the dot, offering sage advice, and cheering me on through the last five years. This is as much your feat as it is mine.

Last but far from the least, to my advisor, Dr. Heather Urry – I feel so incredibly lucky that you decided to take a chance on me five years ago. Thank you for teaching me what it truly means to do good science regardless of the results, for being on board if not more enthusiastic about my medium orchid graphs, for having the same love for reality TV and hatred for musicals, for road tripping to Lancaster, PA and back in the same day, and for unconditionally supporting me throughout these last five years. It's been an honor to learn from you and I would do it all over again if I could.

Portions of this text and ideas presented in this dissertation were published as a mini review of the literature (Patel & Urry, 2024).

Dedication

This dissertation is dedicated to my parents – Ma and Dad – everything I am, I am because of you.

Table of Contents

List of Tables

List of Figures

Discrete and Dimensional Approaches to Affective Forecasting Errors

How will I feel when I try the new restaurant in town? How will I feel if I go to the movies instead of writing my paper? How will I feel when I submit my grant proposal? People make predictions about how they will feel in the future on a daily basis. This process is known as affective forecasting (Wilson & Gilbert, 2003). Most affective forecasting research has focused on the idea that people make affective forecasting errors. Specifically, this research suggests that people tend to inaccurately overestimate the intensity and duration of their future emotions (Gilbert et al., 1998). For instance, people overestimate how nervous they will feel when running a race (Aitken et al., 2021) and how negative they will feel when their preferred candidate loses the presidential election (Barber et al., 2023).

Despite the evidence for affective forecasting errors, specifically those of *overestimation,* some studies also suggest that people may sometimes *underestimate* their affect (e.g., Ruby et al., 2011; Lench et al., 2011; Zelenski et al., 2013), or even make accurate affective forecasts (e.g., Levine et al., 2012; Lench et al., 2019; Moeck et al., 2022). Findings are, thus, inconsistent.

Understanding the reasons for inconsistent findings is crucial for theoretical and practical reasons alike. Theoretically, understanding the sources of inconsistencies can inform the inferences that researchers make from their studies and allow them to make nuanced claims about affective forecasting errors. Practically, affective forecasts are pervasive in people's everyday lives; they guide the situations that people choose to immerse themselves in (Urry & Gross, 2010), are linked to performance on tasks (e.g., Kaplan et al., 2020), and decision-making in domains such as healthcare (e.g., Hoerger et al., 2016) and travel (e.g., Karl et al., 2021). Thus, better understanding the sources of affective forecasting errors can improve our collective understanding of their effects on these downstream processes (Patel & Urry, 2024).

Past researchers have examined two potential sources of variation in findings about affective forecasting errors for the intensity of future emotion (Levine et al., 2012; Mathieu & Gosling, 2012). First, Levine and colleagues (2012) found that when people were asked to imagine an event, make forecasts about how they would feel, and later report how they actually felt *in reference to that event*, they made relatively accurately predictions as opposed to when they were asked to imagine an event, forecast how they would feel, and later report how they felt in general, *without any reference to the event*. Second, Mathieu and Gosling (2012) found through a meta-analysis that when researchers adopted an "absolute" approach (i.e., computed the difference between forecasted and actual affect), people were inaccurate at predicting their emotions, as opposed to when they adopted a "relative" approach (i.e., computed the correlation between forecasted and actual affect), people were fairly accurate (Patel & Urry, 2024).

In this research, I focus on a novel source of variation – the divergent measurement approaches used to measure feelings in affective forecasting studies. Accordingly, I first describe two broad theoretical approaches to emotion research, discrete and dimensional. Arguably, researchers' emotion theories guide their corresponding measurement approaches (Patel & Urry, 2024). I then present four studies – two studies that examine the effect of discrete and dimensional approaches on affective forecasting error and two stimuli validation studies. Finally, I conclude with a general discussion.

I focus on affective forecasts about future emotion intensity since forecasts about intensity have been examined more extensively than forecasts about duration. Additionally, I focus on absolute errors (the difference between predicted and actual affect) as opposed to relative errors (the correlation between predicted and actual affect) since the former analysis technique is more common than the latter. Lastly, since larger scores when computed as the difference between predicted and actual affect indicate larger affective forecasting errors, I used the term affective forecasting “errors” as opposed to affective forecasting “accuracy” to describe this phenomenon throughout this dissertation.

Discrete and Dimensional Approaches to Emotion Research

Broadly speaking, there are two categories of emotion theories. According to the basic/discrete emotions theory, humans have evolved to have a set of basic emotions in response to threats and challenges in their environments (Ekman, 1992; Tooby & Cosmides, 2008). This model proposes three main features of emotions – first, that they have evolved to serve distinct adaptive functions. For example, the emotion of fear is believed to have evolved to help us flee predators and other sources of threat (Öhman & Mineka, 2003). Second, each discrete emotion has its own unique neural signature in the central nervous system that, once activated, leads to its own signature profile of physiology, behavior, and cognition (Posner et al., 2005). Continuing with the example of fear – it activates a specific network of brain regions that leads to a racing heartbeat, increases in skin conductance, widening of the eyes, and other overt behaviors. Third, discrete/basic emotion theorists believe that while people across cultures might interpret emotions slightly differently and even create their own emotion concepts, certain core emotions are innate and thus universal across people and cultures (Ekman & Freisen, 1971). For instance, the expression of the emotion of fear was identified by people from New Guinea who had little to no exposure to Westerners or Western culture (Ekman & Friesen, 1971).

Researchers who adopt a purely discrete emotions approach measure and analyze each emotion as its own category. For instance, in the affective forecasting literature, Aitken and colleagues (2021) asked participants to rate how much excitement, confidence, pride, frustration, and nervousness they expected to experience and actually experienced. Subsequently, they conducted separate paired samples *t*-tests for *each* discrete emotion to examine mean differences in predicted and actual intensity (Patel & Urry, 2024).

By contrast, researchers who adopt the dimensional (or core affect) theory of emotion, conceptualize emotions as combinations of broader underlying processes or dimensions. While there are several two-dimensional models (e.g., positive and negative affect [Watson et al., 1999], approach and withdrawal [Lang et al., 1998]), here I consider the affective circumplex model, comprising dimensions of valence and arousal (Russell, 1980) (see Figure 1). In this model, valence, as displayed on the x-axis, refers to the level of unpleasantness to pleasantness, and arousal, as seen on the y-axis, refers to the level of activation. Accordingly, each emotion is a linear combination of some level of valence and arousal. For instance, fear is an emotion that is conceptualized as a combination of negative valence and high arousal (Posner et al., 2005). Hence, fear is situated in the upper left quadrant of Figure 1, along with other high arousal negative emotions such as anger and frustration. The upper right quadrant comprises emotions that are a combination of high arousal and positive valence such as excitement and elation. The lower half of this circumplex contains the low arousal negative quadrant including emotions such as guilt, and regret, and the low arousal positive quadrant including emotions such as contentment and calmness (Patel & Urry, 2024).

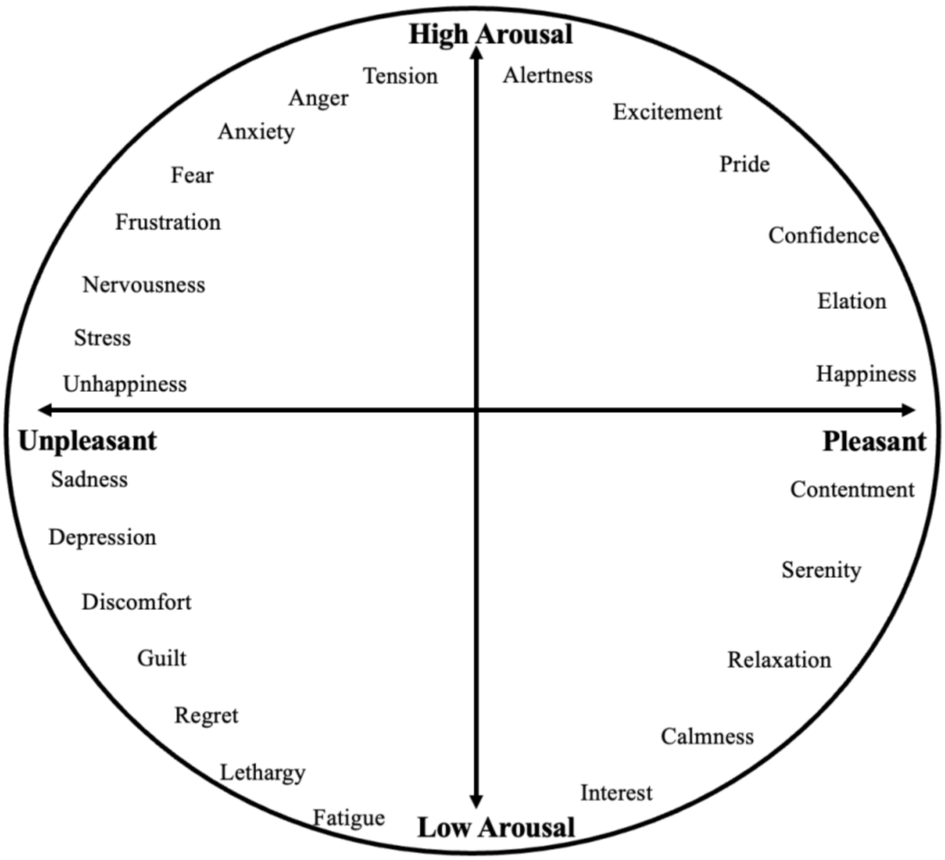

Figure 1. Affective circumplex (adapted from Russell, 1980) displaying valence (x-axis) and arousal (y-axis)

Researchers who adopt a purely dimensional approach may measure and analyze valence and/or arousal directly. For instance, in the affective forecasting literature, Aitken and colleagues (2021) measured predicted and actual valence and arousal using an emoji grid. They then analyzed mean differences in predicted and actual valence and arousal using paired samples *t*-tests (Patel & Urry, 2024).

Lastly, researchers might take a hybrid approach by measuring several discrete emotions and combining them into composite indices of positive and negative affect for analysis. For instance, Barber and colleagues (2023) measured predicted and actual levels of three discrete low arousal positive (calm, relaxed, content) and negative emotions (bored, lonely, sluggish), and three discrete high arousal positive (excited, enthusiastic, activated) and negative emotions (angry, anxious/worried, disappointed). Subsequently, they combined these discrete emotions into composite indices of low arousal positive affect, low arousal negative affect, high arousal positive affect, and high arousal negative affect, respectively, in their statistical analyses. When researchers use this hybrid approach (i.e., measure discrete emotions but analyze dimensional affective states), their participants are ultimately forecasting and rating *discrete* emotions. Thus, even though they ultimately combine these discrete emotions into dimensional affective states and analyze those, participants made ratings about discrete emotions and presumably the processes underlying their ratings are those that occur when they rate discrete, not dimensional states.

Reasons Why Affective Forecasting Errors May Differ Based on Measurement Approach

In the following section, I discuss three main reasons to believe that affective forecasting errors could vary based on measurement approach: 1) discrete emotions and dimensional states are distinct, 2) the processes related to making predictions about discrete emotions and dimensional states could be different, and 3) the processes related to rating actual discrete emotions and dimensional states could be different.

First, since discrete emotions and dimensional affective states are theoretically distinct, people might accordingly be differentially accurate at predicting them. Forecasting errors could be larger for dimensional states like valence and arousal since, according to discrete emotion researchers, these states are more abstract and nebulous, as compared to discrete emotions like fear that have universal, well-defined characteristics (Ekman, 1992). Alternatively, forecasting errors could be smaller for valence and arousal since, according to dimensional researchers, these states represent core affective processes that underlie the experience of any emotion (Barrett, 1998). Furthermore, people may find it easier to make predictions about the intensity of these core affective processes, as opposed to identifying, labeling, and predicting the intensity of individual discrete emotions (Patel & Urry, 2024).

Second, it is possible that the processes that occur when affective forecasting errors are made about discrete emotions are different from those that occur when affective forecasting errors are made about dimensional states. Since affective forecasting errors comprise of two components - affective forecasts (predicted affect/emotions) and actual affect/emotions, people may have different affective forecasting errors because of differences in the processes underlying predicted affect, actual affect, or both (assuming that differences between predicted and actual affect do not cancel each other out).

According to past research, the process of affective forecasting comprises three steps - first, people create mental simulations or “previews” of future events. Second, their previews induce hedonic reactions, or “premotions” in the present. Third, people then rely on the contexts that they are currently in and their simulations and premotions to create affective forecasts (Gilbert & Wilson, 2007; 2009). Discrete researchers could argue, for example, that people may be able to simulate situations involving discrete emotions more vividly than those involving positive/negative affect. This could also mean that they experience stronger premotions, and as long as their premotions accurately reflect the target situation, this might lead to smaller affective forecasting errors than if researchers were to use a dimensional approach. Alternatively, dimensional researchers could argue that people may be able to easily simulate situations involving overall general feelings of positive/negative affect, rather than those involving specific discrete emotions. This could, in turn, induce stronger premotions, and lead to smaller affective forecasting errors, than if researchers were to use a discrete approach (Patel & Urry, 2024).

On the other hand, according to the Modal Model of Emotion (Gross & Thompson, 2007), the process of experiencing actual emotions involves a situation - attention - appraisal - response sequence (Gross & Thompson, 2007; Gross et al., 2011). People experience emotions when they are in a situation that has some psychological self-relevance or importance. They pay attention to this situation or aspects of it, evaluate or "appraise" what the situation means with respect to their goals, and accordingly experience an emotional response (Gross & Thompson, 2007). Since discrete emotions are typically elicited by specific situations (Gross & Thompson, 2007), discrete researchers could accordingly argue that people are more likely to attend to them, appraise them and experience stronger emotions than dimensional affective states that do not necessarily require external stimuli. This could in turn, lead to differences in affective forecasting errors for discrete versus dimensional states. However, dimensional researchers might argue that core affect underlies all emotional experiences, is the "heart of any emotional episode" (Russell & Barrett, 1999, p. 806), and that a person is constantly experiencing some form of core affect, even if it is a neutral state (Russell & Barrett, 1999). As such, dimensional researchers could suggest that people might be more attuned to fluctuations in their core affect because it reflects a constant, dynamic state of being, rather than requiring a discrete situational trigger. This could, in turn, lead to differences in affective forecasting errors for discrete versus dimensional states.

There is limited evidence for affective forecasting errors about the intensity of emotions/affect differing based on measurement approach, in part because researchers rarely use *both* approaches in one single study. However, Aitken and colleagues (2021) used both approaches in one study to examine forecasts about *running* in a trail race and *crossing* the finish line. They found that people underestimated their confidence and overestimated their nervousness with respect to running in the race. Additionally, they underestimated their confidence and pride but overestimated their frustration regarding crossing the finish line. People were accurate at predicting their valence when crossing the finish line; however, they overestimated their arousal. These findings suggest that affective forecasting errors may differ based on measurement approach. However, this study did not directly answer the current question of interest, since its goal was not to directly compare affective forecasting errors for discrete versus dimensional states. Overall, due to the limited evidence from just one study, it is currently unclear whether affective forecasting errors differ when measured and analyzed using a discrete versus a dimensional approach.

Present Studies

To date, most of the affective forecasting research has used either a discrete or a dimensional approach to measuring and analyzing affective forecasting errors (Patel & Urry, 2024). However, none of these studies have utilized both approaches in a single study to compare affective forecasting errors. Furthermore, most studies have examined affective forecasting errors about events such as presidential elections (Dunn et al., 2007) and football games (Wilson et al., 2000) that likely induce high arousal emotions (e.g., anger). Less common are studies that examine affective forecasts about mundane events such as completing tasks at work (Kaplan et al., 2020) that likely induce lower arousal emotions (e.g., boredom). However, given that people typically experience events that likely induce both low and high arousal emotions in their daily lives, it is crucial to examine affective forecasting errors for low and high arousal positive and negative emotions alike.

The goal of the present studies is to fill these gaps in existing research and advance our understanding of affective forecasting errors. The studies will do so in two ways. First, they will examine whether forecasting errors differ for the intensity of discrete emotions versus dimensional affective states. Second, they will examine discrete versus dimensional affective forecasting errors for a variety of low and high arousal positive and negative emotions in response to low and high arousal stimuli.

In Study 1, I examined affective forecasting errors in a within-subjects study. On the first day of the study, participants saw descriptions of pictures and made predictions about how much they would feel various discrete emotions and dimensional affective states when they viewed these pictures one week later. One week later, they viewed the pictures that they had seen descriptions of and rated how much they actually felt the same discrete emotions and dimensional affective states. I used these ratings to examine whether affective forecasting errors differed for discrete versus dimensional states.

In Study 3, I examined affective forecasting errors using a between-subjects study. I used pictures and picture descriptions that were validated in Studies 2a and 2b to examine affective forecasting errors for high and low arousal positive and negative emotions. I asked participants to make predictions about how much they would feel either discrete *or* dimensional states based on descriptions of pictures, report how they actually felt when viewing them, and compared affective forecasting errors using both measurement approaches.

Overall, this set of studies is poised to provide a holistic understanding of affective forecasting errors since it collectively examined discrete emotions and dimensional affective states that spanned the entire affective circumplex. Furthermore, these studies used a combination of tightly controlled within-subjects and between-subjects experimental designs. This enabled us to provide an in-depth, nuanced understanding of affective forecasting errors, which were previously thought to be "all-bad" or overall "error prone."

Transparency Statement

Data, materials (Studies 2a, 2b, and 3) and analysis code (all studies) for this dissertation are available in the Open Science Framework project at https://osf.io/fsnu9. I preregistered the hypothesis, method, and data analysis plan for Study 3 prior to data collection on OSF at https://osf.io/shv3x. I report how I determined my sample sizes, all data exclusions, all manipulations, and all measures in the studies.

Study 1

The purpose of this study was to examine whether affective forecasting errors differed when they were measured using discrete and dimensional approaches. I aimed to examine these errors for a variety of picture-induced discrete emotions and dimensional states that spanned the affective circumplex.

I tested these ideas via secondary analysis of existing data collected as part of a larger study that ran for 9 consecutive days (Petagna & Wormwood, 2023). On the first day of the study (Day 1), participants completed an affective forecasting task. They were told that they would view 20 pictures in one week, but during the current session they would view descriptions of each of those pictures and make predictions about how they would feel when they actually viewed the picture in the future. On the last day of the study (Day 9), participants completed a picture viewing and affect rating task. In this task, they viewed the pictures described previously and rated how they actually felt using the same scales as in the previous task.

The researchers also utilized an experience sampling protocol on Days 2-8 (i.e., between the affective forecasting task and picture viewing and affect rating task) and collected data on various constructs and at multiple points through the days. I will only focus on Days 1 and 9 since these were the days on which participants made their affective forecasts and actual affect ratings. More information on the constructs and results of the experience sampling part of the study (Days 2-8) can be found in Petagna and Wormwood (2023).

My primary research question was as follows: Do affective forecasting errors differ when measuring discrete emotions versus dimensional affective states? I was able to answer this question as Petagna and Wormwood (2023) measured affective forecasting errors for discrete and dimensional states. Analytically, I was able to do so by including measurement type (discrete, dimensional) and type of affect ratings (predicted, actual) as factors in my model.

If I found that affective forecasting errors were different for discrete emotions than for dimensional affective states, this would suggest that some of the variation in existing findings might be accounted for by measurement approaches. Additionally, this would indicate that processes involved in predicting and/or generating discrete emotions are somehow different than the processes involved in predicting and/or generating dimensional affective states. Future research would be needed to understand those differences. In contrast, if I found that affective forecasting errors did not differ for discrete emotions compared to affective dimensional states, this would indicate that something other than variation in measurement approaches explains variation in affective forecasting errors across studies. Future research would be needed to identify other sources of variation. Alternatively, null findings could also suggest that aspects of study design were obscuring any true differences between affective forecasting errors for discrete versus dimensional approaches. I did not have *a priori* predictions about which approach would show larger errors than the other nor the direction of those errors.

Method

Design

The study used a within-subjects design and an experience sampling protocol. There were two predictors of interest. The predictor variables included type of affect rating (predicted, actual) and measurement type (discrete, dimensional). Both predictors were manipulated within-subjects. The dependent variable was the level of affect rating.

Participants

Based on an *a priori* power analysis, the original researchers determined that they required 105 participants to detect associations of medium effect size for their original study (Petagna and Wormwood, 2023; Tabachnick & Fidell, 2018). Note that this power analysis was conducted for associations that are not relevant to this current study. To account for attrition, they aimed to recruit participants until they had a sample of 150 participants or until the end of the academic semester.

Participants were recruited from the University of New Hampshire. To be eligible to participate in the study, participants had to be over 18, fluent in English, and have access to their own smartphone. Participants were compensated with up to 2 credit hours in an undergraduate psychology course: 1 credit hour for completing the surveys at the beginning and end of the study, and 1 credit hour for completing at least 42 out of 56 experience sampling surveys over a 1-week period.

The consent survey on Day 1 of the study was completed by 186 participants. Of these participants, 161 began the experience sampling portion of the study, and 131 completed the final survey at the end of the study. Of these 131 participants, the researchers excluded five participants: two were removed prior to analyses for not completing enough experience sampling prompts (< 12 prompts, < -3 *SD* from mean), and three participants were removed from the sample as outliers (± 3 SD from the mean) on various measures during analysis.

The final sample (N = 126) was 18-31 years (M = 19.22, SD = 1.96). A majority of participants were non-Hispanic and white (97.6%), with the remaining participants identifying as one or more of the following: Hispanic/Latino, East Asian, Native Hawaiian or other Pacific Islander, Indigenous, Native American, or Alaska Native, or other category not provided. The sample also comprised participants who self-identified as female/woman (75.4%), male/man (23%), and non-binary or agender (1.6%).

Materials

Picture Stimuli. Twenty emotionally evocative pictures were taken from the International Affective Picture System (Lang et al., 2008). Pictures were selected based on normed ratings of their valence (pleasantness/unpleasantness) and arousal (activation/deactivation) such that five pictures were drawn from each of the four combinations of valence and arousal (high arousal positive, high arousal negative, low arousal positive, low arousal negative). Examples included a person skiing (high arousal positive), a snake hissing (high arousal negative), a child smiling (low arousal positive), and a woman sitting alone at a restaurant (low arousal negative). Petagna and Wormwood (2023) also conducted manipulation check analyses that demonstrated the expected pattern of ratings for predicted and experienced valence and arousal.

Picture Descriptions. For each of the pictures, the original authors developed a brief description to provide to participants. Descriptions were 8 to 18 words long (M = 12.40, SD = 3.23). Example descriptions included: "A skier at the top of a steep long-distance ski jump" and "A snake hanging off a branch posed to strike with its jaws open." All descriptions are provided in the Supplementary Materials.

The original authors validated these picture descriptions in a separate sample (N = 115) who were between the ages of 18 to 83 and recruited via Prolific (M_{age} = 38.37, SD_{age} = 13.72, 79.1% white; 58.3% female/woman). These participants were asked to read each description and imagine the corresponding image. Participants then viewed the picture and rated how similar it was to what they imagined on a 7-point scale from 1 (*not at all similar*) to 7 (*extremely similar*). In general, participants found the pictures quite similar to what they imagined (*Grand Mean* = 4.99, SD = .59); although, there was some variation across descriptions (range: 3.75 – 5.67).

Tasks and Measures

Affective Forecasting Task. During Part 1 of the study, participants were informed they would view 20 emotionally evocative pictures in one week, and that they would be presented with a brief description of each picture during the current study session. Participants were shown one picture description at a time in a random order. After reading each description (untimed), they were asked to "predict how much you think you will feel <emotion/affective state> when viewing the picture in the future" on a 1 (*Not at all*) to 7 (*Very much*) scale. Participants were asked to make predictions about 16 discrete emotions including happy, excited, relaxed, focused, content, tired, neutral, sad, nervous, frustrated, stressed, annoyed, proud, grateful, angry, and afraid. The original researchers selected 12 of these emotions based on how often people in a previous experience sampling study freely generated them when describing their current emotions (Hoemann et al., 2020). They additionally added four high arousal emotions (two positive: grateful, proud; two negative: angry, afraid), as these emotions were not represented in the previous selection.

Participants were then also asked to predict their general affective reaction (valence and arousal) to each picture. Specifically, they were asked to "please rate how pleasant or unpleasant you think you will feel while viewing the image described" on a 1 (*unpleasant*) to 7 (*pleasant*) scale. They were also asked to "please rate how activated or deactivated you think you will feel in your body while viewing the image described" on a 1 (*deactivated*) to 7 (*activated*) scale. Participants made these valence and arousal ratings after viewing descriptions for each picture after they rated the discrete emotions.

Picture Viewing and Rating Task. One week later, participants viewed the 20 pictures described in the affective forecasting task one at a time in a random order, and were asked to "rate how much <adjective> you are currently experiencing" on the same 1 (*Not at all*) to 7 (*Very much*) scale. Additionally, they rated their current valence and arousal after viewing each picture using the same scales for valence [1 (*unpleasant*) to 7 (*pleasant*) scale] and arousal [1 (*deactivated*) to 7 (*activated*)] as in the affective forecasting task.

Affective Forecasting Error. The original authors of this research (Petagna and Wormwood, 2023) uploaded their aggregated data and materials for the larger study on the Open Science Framework (OSF) website at: https://osf.io/5berd/. Upon request, Petagna and Wormwood (2023) shared their raw, trial-level data with me for this study.

Each participant had affective forecasting ratings for 15 discrete emotions (I excluded neutral) and 2 dimensional states (valence and arousal) for each of the 20 picture descriptions. Thus, they each made 340 predicted affect ratings. Similarly, each participant also had actual affect ratings for the same 15 discrete emotions and 2 dimensional states for each of the 20 pictures. Thus, they each made 340 actual affect ratings. Overall, each participant (except for one) had 680 observations that were included in the confirmatory analyses. One participant had 340 observations - only affective forecasts - and no ratings of actual affect. Thus, the confirmatory analysis included a total of 85340 observations.

Procedure

The study ran in three parts over nine consecutive days. On Day 1, participants completed Part 1, an online session via Qualtrics lasting approximately 45 minutes in which they provided consent and completed the Affective Forecasting Task and other questionnaires. They were also given instructions on how to enroll in Part 2, the experience sampling part which lasted for seven days. Part 2 is irrelevant for the current investigation since I did not use any of the measures and data from it. However, more details about this part can be found in Petagna and Wormwood (2023). The day after participants completed Part 2 of the study, they received a final notification asking them to complete Part 3, the Picture Viewing and Affect Rating Task, which they completed via Qualtrics on their smartphone. Please see Figure 2 for a schematic of the procedures.

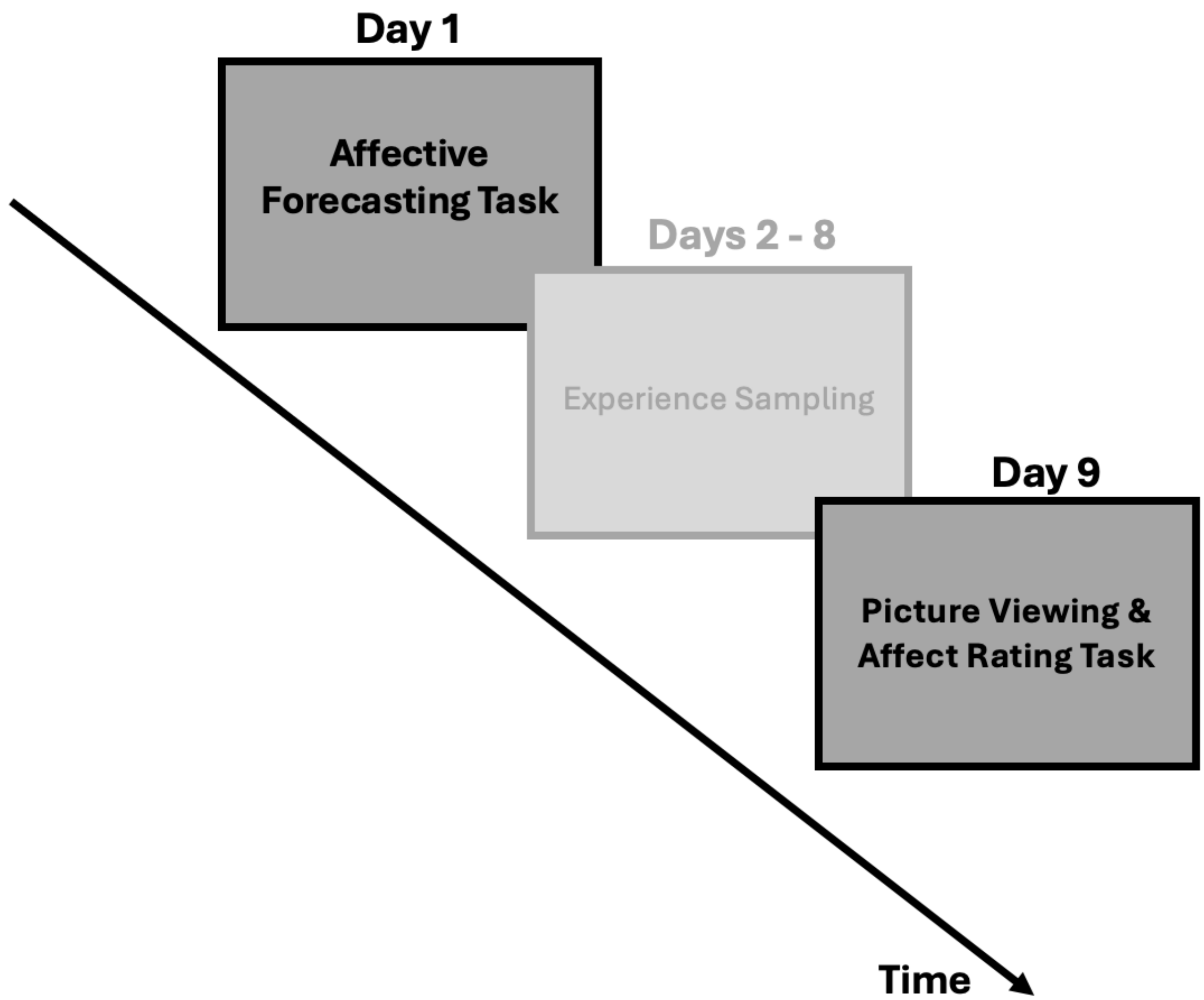

Figure 2. Schematic of Study 1 procedure. The grayed-out parts of the schematic are days for which I did not analyze data. The bolded parts of the schematic are days for which I analyzed data for my primary research question in this study

Data Analytic Strategy

To test whether discrete and dimensional measurement approaches predicted unique patterns of affective forecasting errors, I conducted a linear mixed effects regression examining sources of variation in affect ratings. I used *R* (Version 4.4.1) to conduct this model with the *lme4* package (Bates et al., 2015).

The linear mixed effects model included fixed effects for type of affect ratings, measurement type, and their interaction, as well as random intercepts for participants, emotions, and stimulus numbers. The emotions variable captured which emotion participants were making a rating about (e.g., anger, sadness, valence, arousal and so on). The stimulus number variable captured which picture description/picture combination participants were rating (e.g., picture # 1 and its description, picture #2 and its description, and so on). Including the random intercepts for participants, emotions, and stimulus numbers would allow the model to detect variation in the data that could be accounted for by participants having multiple ratings, emotion items being different from one another, and stimuli being different from one another. Lastly, I deviation-coded type of affect ratings (predicted affect ratings = 0.5, actual affect ratings = -0.5) and measurement type (dimensional = 0.5, discrete = -0.5).

I first ran an empty model without the fixed effects to examine the percentage of the total variance in affect ratings that was explained by participants, emotions, and stimulus number. I found that 6.2% of the variation was explained by differences among participants in the sample. Additionally, 11.9% of the variation in ratings was explained by differences among emotions. Lastly, 2.4% of the variation in ratings was explained by differences among stimuli (i.e., pictures and their descriptions). Collectively, this suggested that there was clustering in the data and that utilizing multilevel modeling was an appropriate analysis method. Therefore, I retained the random intercepts for participants, emotions, and stimulus numbers in my confirmatory model.[1,2]

I also conducted equivalence tests to examine whether effects fell inside a region of practical equivalence (ROPE). The ROPE for this test [-0.21, 0.21] was determined in R based on default guidelines tied to the scale of the data. If the effect was within the ROPE, it was small enough to be negligible or equivalent to 0. I also report second-generation *p*-values (SGPV), that indicate the proportion of the confidence interval that falls within the ROPE. Values closer to 1 provide strong evidence that the effect is essentially equivalent to 0.

Results

Confirmatory Analysis

Overall, as seen in Table 1, the fixed-effect predictor variables (type of affect rating, measurement type) collectively accounted for 7% of the variance in affect ratings, the dependent variable. There was a significant fixed effect of type of measurement type, suggesting that participants (regardless of whether they were rating predicted or actual affect) rated dimensional states (M = 4.30, *95% CI* [3.60, 4.90) as significantly more intense than discrete emotions (M = 2.60, *95% CI* [2.30, 2.80]), B = 0.17, z = 4.90, p < .001.

Additionally, there was no significant fixed effect of type of affect rating, suggesting that participants (regardless of which emotion/affective state they were rating) predicted (M = 3.40,

---

[1] The analysis was conducted using the *lme4* package (Bates et al., 2015) in R and utilized the following structure:

```
lmer(rating ~ affect_type*measure_type + (1 |ID) + (1|emotion) + (1 |stimNum),
        dataframe,
        contrasts = list(affect_type = affect_type.deviation.con,
                         measure_type = measure_type.deviation.con))
```

[2] I included a by-participant random slope for type of affect rating in the confirmatory model, which suggested that there were some individual differences in affective forecasting errors. However, conclusions regarding the confirmatory hypothesis did not change (i.e., the hypothesized two-way interaction between type of affect rating and measurement type was not statistically significant). Thus, I reported the model without this by-participant random slope in the main text.

Table 1

Results from linear mixed effects regression on affect ratings, including fixed effects for type of affect rating, measurement type (both deviation-coded), and random intercepts for participants, emotions, and stimulus number

	Affect Rating		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	3.41	3.03 – 3.78	<0.001***
Predicted-Actual	0.02	-0.02 – 0.06	0.373
Dimensional-Discrete	1.70	1.02 – 2.37	<0.001***
Predicted-Actual × Dimensional-Discrete	0.05	-0.03 – 0.13	0.203
Random Effects			
σ^2	3.42		
$\tau_{00\ ID}$	0.27		
$\tau_{00\ stimNum}$	0.10		
$\tau_{00\ emotion}$	0.21		
ICC	0.14		
N_{ID}	126		
$N_{emotion}$	17		
$N_{stimNum}$	20		
Observations	85340		
Marginal R^2 / Conditional R^2	0.070 / 0.204		

Note. *p < .05; **p <.01; ***p < .001

95% CI [3.04, 3.80]) and actually experienced (M = 3.40, *95% CI* [3.02, 3.80]), B = 0.02, z = 0.89, p = .370, similar levels of emotions or affect. Furthermore, results from the equivalence test indicated that the fixed effect of type of affect rating was small enough to be equivalent to zero, with a 95% CI for the effect size falling within the ROPE [-0.21, 0.21], *95% CI* [-0.01, 0.05], and a statistically significant *p*-value of .001. The SGPV was >0.999, indicating strong evidence for equivalence. Thus, equivalence was accepted.

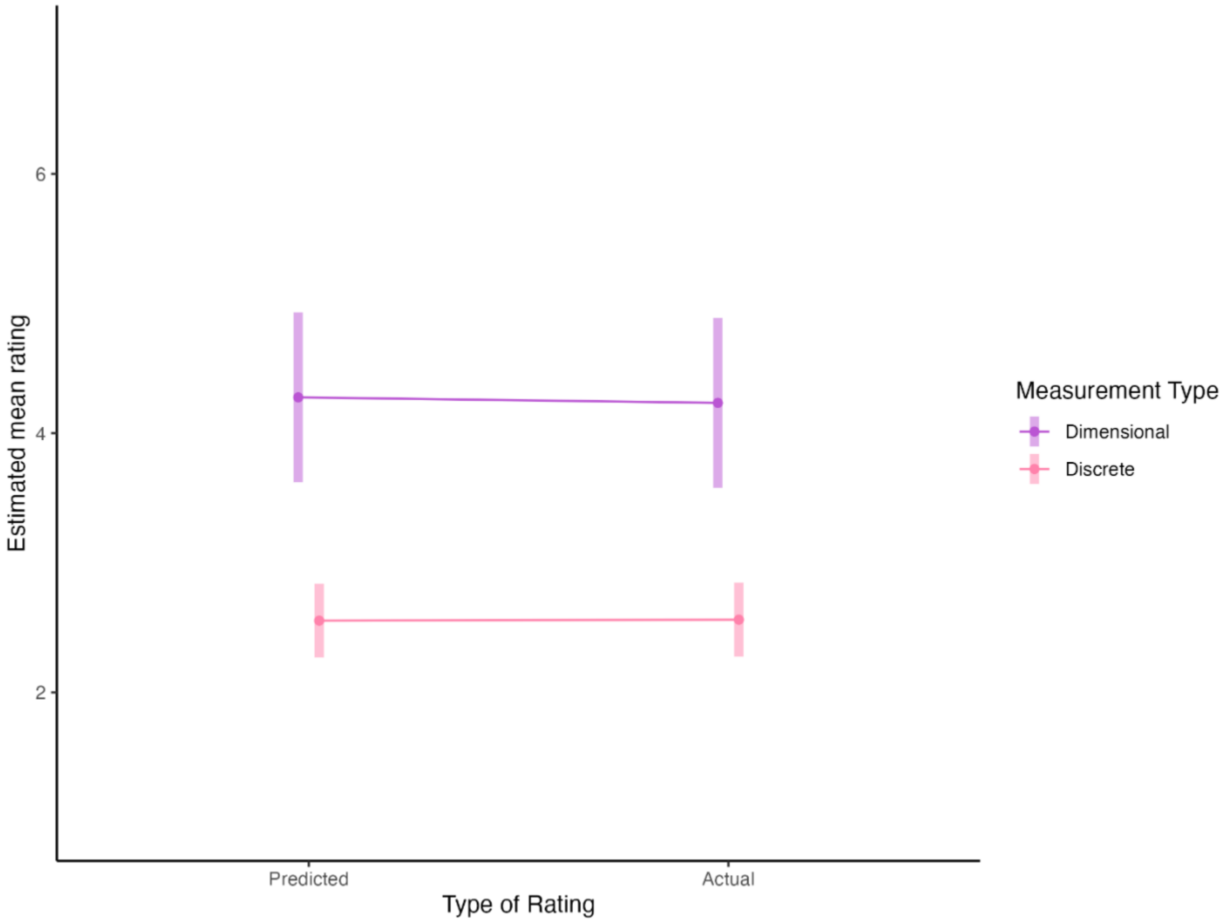

Figure 3. Estimated mean affect ratings as a function of measurement type (dimensional, discrete) and type of affect rating (predicted, actual). Error bars represent 95% confidence intervals around each mean, shown in filled circles.

The hypothesized two-way interaction between type of affect rating and measurement type was not statistically significant (see Figure 3). Thus, contrary to our confirmatory hypothesis, participants exhibited similar affective forecasting errors for discrete and dimensional states. This interaction between type of affect rating and measurement type was small enough to be equivalent to zero, with a 95% CI for the effect size falling within the ROPE [-0.21, 0.21], *95% CI* [-0.01, 0.11], and a statistically significant *p*-value <.001. The SGPV was >0.999, indicating strong evidence for equivalence. Thus, equivalence was accepted.

Exploratory Analysis

The goal of Study 1 was to examine directional affective forecasting errors such as underestimation and overestimation through raw affect rating scores. However, given that Petagna & Wormwood's original study (2023) and other affective forecasting researchers' (e.g., Frank et al., 2021) studies have often focused on the magnitude of affective forecasting errors through the absolute value of affective forecasting error scores, I also applied the confirmatory model on absolute affective forecasting scores (i.e., |predicted-actual affect|). Consistent with the confirmatory analysis, I found no statistically significant effect of measurement type on absolute affective forecasting error scores. However, I did find that there were overall affective forecasting errors of about 1 unit ($B_0 = 0.94$, *95% CI* [0.79, 1.09], $p < .001$). Please see Table S1 in the Supplementary Materials for the results of this model. Furthermore, unlike the confirmatory analysis on directional affective forecasting errors, results from the equivalence test indicated that the absolute affective forecasting errors were not equivalent to zero, with a 95% CI for the effect size falling outside the ROPE [-0.10, 0.10], *95% CI* [0.81, 1.06], and a non-significant *p*-value of >.999. The SGPV was 0.764, indicating weak evidence for equivalence.

Furthermore, given that Petagna and Wormwood (2023) showed participants picture stimuli that varied in terms of their levels of valence and arousal, I also explored the confirmatory model by including fixed effects of stimulus valence (positive, negative) and stimulus arousal (high, low) in addition to type of affect rating, measurement type, and the interactions between all four fixed effects. As in the previous model, I also included random intercepts for participants, emotions, and stimulus numbers. Furthermore, I deviation-coded stimulus valence (pleasant = 0.5, unpleasant = -0.5), and stimulus arousal (low = 0.5, high = -0.5). Overall, I did not find a statistically significant four-way interaction between stimulus valence, stimulus arousal, type of affect rating, and measurement type, or a statistically significant three-way interaction between stimulus valence or stimulus arousal, type of affect rating, and measurement type. Please see Table S2 in the Supplementary Materials for the results of this model.

Lastly, I also explored the confirmatory model using a hybrid approach for discrete emotions. Specifically, I created an overall score for discrete positive emotions by averaging scores on 7 positive discrete emotions per picture and an overall score for discrete negative emotions by averaging scores on 8 negative discrete emotions per picture. Additionally, I had scores for valence and arousal per picture already. I repeated the confirmatory analysis and included a fixed effect for emotion valence (deviation-coded) that captured the overall scores for positive discrete, negative discrete, valence and arousal. I removed the fixed effect for measurement type (since this information was captured in the emotion valence variable). I also remove the random intercept for emotions since this was no longer needed. Consistent with the confirmatory analysis, I found no statistically significant interaction effect between type of affect rating and emotion valence. Please see Table S3 in the Supplementary Materials for the results of this model.

Discussion

The goal of Study 1 was to examine whether affective forecasting errors varied when measured using a discrete versus a dimensional approach. Contrary to the hypothesis, affective forecasting errors did not differ for discrete emotions versus dimensional affective states.

Additionally, participants did not make *directional* affective forecasting errors, but did exhibit affective forecasting errors in *absolute* terms.

In this study, I found evidence of insufficient variation in affective forecasting errors. This was true in both the raw and absolute data. Insufficient variation was apparent in the fact that most values for the predicted - actual differences were near zero. Thus, it is possible that the restricted range of affective forecasting error scores made this study ill-suited to finding variation in these scores due to measurement type.

The contradictory findings that participants did not make *directional* affective forecasting errors, but did exhibit affective forecasting errors in *absolute* terms might seem perplexing at first. However, it is likely that some people underestimated their affect while others overestimated it, and these directional errors cancelled each other out in the raw, trial-level data.

One potential reason that could explain the restricted range of forecasting errors and absolute affective forecasting error scores was the fact that Petagna and Wormwood (2023) conducted their study on smartphones. It is possible that viewing pictures on the small screen of a smartphone may have reduced the intensity of affect that people expected to experience and actually experienced (the intercept or average amount of predicted and experienced emotion/affect intensity was $B_0 = 3.41$ on a 1-7 scale). Overall, had I found greater variation in affective forecasting errors, I might have been better positioned to detect any changes in these errors based on measurement type.

Despite the restriction of range in affective forecasting errors, this study has several strengths that are worth noting. First, it used a within-subjects design and had participants use *both* discrete and dimensional measures of affective forecasting errors in a single study. Due to the within-subjects design, variability between participants was minimized and this increased the chances of detecting a true effect. Second, Petagna and Wormwood (2023) also examined forecasts for a variety of emotions and affective states, unlike past research in this area that has examined affective forecasts for a limited number of emotions/affective states. Lastly, Petagna and Wormwood (2023) asked participants to make their affective forecast and affect ratings one week apart. This likely reduced the chances of participants recalling their affective forecasts at the time that they made their actual emotion/affect ratings after viewing the pictures, and presumably prevented their actual emotion/affect ratings from being influenced by their forecasts.

Aside from the strengths of this study, there are a variety of limitations that could explain why I may not have found a difference in affective forecasting errors for discrete emotions versus dimensional affective states. Before unpacking the limitations of this study, it is worth mentioning that the goal of the original study by Petagna and Wormwood (2023) was *not* to compare affective forecasting errors for discrete versus dimensional measures. Rather, these limitations speak more to the *way I used their data* to achieve the present aims than to the original study.

While one of the strengths of this study was certainly that it used both discrete and dimensional approaches to test affective forecasting errors, participants made both discrete and dimensional ratings one after the other, within a relatively short time frame. Since there was no interval between ratings made using discrete and dimensional measures, this could have inadvertently muddied the processes, like the mental simulations and premotions, underlying discrete and dimensional affective forecasting. For example, it is possible that when participants were asked to predict their future happiness, they then simulated a situation meeting the description of the picture, which induced a premotion that aligned with happiness. They may then have used their context, simulation, and premotion to predict their future happiness. Within a span of seconds, participants were then asked to predict, say, their future arousal. It is possible that the simulation and premotion for arousal predictions is influenced by that of happiness forecasts. If this is the case, this could explain why I was unable to detect differences between discrete and dimensional affective forecasting errors. Future researchers could benefit from utilizing a between-subjects design to ensure that there is no carryover between simulations and premotions for discrete versus dimensional forecasts.

Another limitation of this work was that the picture stimuli were normed using a dimensional approach, but ultimately participants rated multiple discrete emotions and two dimensional states for every picture in the study. In particular, the authors selected five photos that induced overall high arousal positive affect, five for high arousal negative affect, five for low arousal positive affect, and five for low arousal negative affect. This norming procedure was dimensional in that it enabled the original authors to induce the expectation and experience of overall dimensional states, but not specific discrete emotions. For example, from a participants' perspective, viewing a picture normed to induce high arousal negative affect, could induce high levels of fear, anger, or both. Given that the original authors ultimately measured *several positive and negative discrete emotions* for each picture, it is possible that some of the emotions that participants rated may not have been applicable to or very highly rated for each image. This could have prevented me from ultimately detecting differences based on measurement approach. Had the original authors normed their stimuli to induce *one specific discrete emotion* and subsequently tested *only that* discrete emotion, this might have positioned us better to detect any differences between affective forecasting errors as measured by both discrete and dimensional approaches.

Additionally, since the original authors recruited a sample primarily comprising students from the University of New Hampshire between the ages of 18 and 31 years, this limits the generalizability of the findings to young adults in the U.S. who are currently enrolled in college.

Past research has found that there are age differences in affective forecasting errors (e.g., Barber et al., 2023; Scheibe et al., 2011). Specifically, in one study, older adults were more accurate at predicting their emotions than younger adults following an election victory, but younger adults were better at predicting their emotions than older adults following an election loss (Scheibe et al., 2011). Thus, findings may only generalize to a very specific sample, and not to all adults across the U.S. Future researchers should aim to diversify their samples by including adults of all ages across the U.S. to increase the generalizability of findings.

Lastly, it is possible that people are not differentially accurate at predicting discrete versus dimensional states. However, these findings are from just one study that used one methodological approach (within-subjects experiment), one conceptualization of discrete emotions (measuring several discrete emotions) and dimensional states (measuring valence and arousal), and has several limitations as noted above. To confidently claim that affective forecasting errors do or do not differ when they are measured using discrete versus dimensional approaches, future research that employs alternative methodological approaches and a different design to test this idea is needed.

To meet this need, I conducted Study 3. In Study 3, I used a between-subjects experiment to examine whether forecasts were different for discrete versus dimensional states regarding picture descriptions and pictures that were validated to induce primarily one specific discrete emotion. Before describing Study 3, I will first describe the two studies (Study 2a and Study 2b) I conducted to validate the stimuli that I used in Study 3.

Study 2a

The goal of this study was to validate the pictures that would be used in Study 3 to elicit specific discrete emotions as the target context in which to examine affective forecasting errors.

Through this study, I aimed to end up with a set of 12 pictures each of which induced *primarily one discrete emotion* with three discrete emotions from each quadrant of the circumplex.

Method

Participants

I recruited 52 adult participants (ages 18 years and up) through *Prolific* (https://www.prolific.co). I determined my sample size based on existing financial resources. I only recruited participants who reported U.S. nationality, were residing in the United States, were fluent in English, had a minimum "approval rate" of 95% on Prolific, had not participated in our lab's previous affective forecasting studies, and were using a laptop/desktop (not a mobile device) to complete the study. Participants were compensated $5 USD for their participation.

One participant reported that they had previously participated in the study and one participant reported using a mobile device. These participants were excluded. Thus, I had a final sample of $N = 50$. Please see Table S4 in the Supplementary Materials for the sample's demographic information. All study procedures (and the procedures for Study 2b and 3) were approved by the Social, Behavioral, and Educational Institutional Review Board at Tufts University prior to data collection.

Materials

Pictures. I selected 36 pictures from a variety of sources: 8 from the International Affective Picture System (IAPS; Lang et al., 2008), 16 from the Nencki Affective Picture System (NAPS; Marchewka et al., 2013) and 12 that were not protected by copyright from Google. All pictures were the same size (1024 x 768 pixels). In case the original picture dimensions were smaller than 1024 x 768 pixels, I added black borders to ensure that they were all the same size. Please see Table S5 in the Supplementary Materials for more information about the picture stimuli. I selected these pictures such that each one would induce anger, sadness, fear, happiness, disgust, fatigue, alert, calm, excited, relaxed, contentment, or boredom.

I selected specific emotions as my goal was to match each discrete emotion to a single picture in the picture task, since I wanted participants to make ratings about only the dominant emotion elicited by that image. Additionally, I selected these emotions such that they represented each quadrant on the circumplex. Specifically, I first selected the basic emotions (anger, sadness, happiness, fear, disgust) (Ekman, 1992) minus surprise since it is neither positively nor negatively valenced. I then used Russell's (1980) affective circumplex model to fill in the gaps and selected additional emotions that would span the entire affective circumplex such that I had three discrete emotions from each quadrant.

Self-report Measures

Primary Experienced Emotion. After participants viewed each picture, they were asked to "select the primary emotion that you experienced when viewing the picture" from a list of 14 options including anger, sadness, fear, happiness, disgust, fatigue, alert, calm, excited, relaxed, contented, boredom, other (please specify), and I didn't experience any emotion.

Secondary Experienced Emotions. After participants answered the question about their primary experienced emotion, they were asked "aside from the primary emotion, what, if any, other emotions did you experience when viewing the picture?" from the same list of 14 response options as before.

Overall Reflection. At the end of the study, participants reported how the study went overall and whether there were any points of confusion or difficulty that affected their participation. They were provided with a text box to type in their answers. If everything went smoothly, they were asked to type "no issues" or something similar. If they did not wish to respond, they were asked to type "decline." The goal of this question was to: 1) find any points of confusion or hiccups in the study and troubleshoot them, and 2) check if participants were reading the questions and participating in good faith. I did not systematically code these responses; however, I examined them and confirmed that the above objectives were met.

Procedure

All study procedures were conducted online via Qualtrics. At the onset of the study, participants provided informed consent and answered eligibility questions. Participants were told that they were about to view a series of pictures and answer questions about their emotions. Participants then saw each picture for 12 seconds. After viewing each picture, they reported their primary and secondary emotions when viewing the picture. This picture - rating cycle occurred 36 times. Lastly, participants answered a demographics questionnaire, the overall reflection question, and viewed a debriefing form.

Data Analytic Strategy

Primary Experienced Emotion. I identified the modal primary experienced emotion for each picture. If the modal primary experienced emotion for each picture was in fact the emotion that I intended to induce (e.g., mode = disgust for the picture intended to induce disgust), I proceeded with using this picture in Study 3. However, if a picture was bi-modal (e.g., mode 1 = disgust and mode 2 = sadness for the picture intended to induce disgust), I planned to exclude that picture from Study 3. In cases where more than one picture's modal primary experienced emotion was the emotion that I intended to induce, I relied on the picture with the highest mode and included that picture in Study 3.

Secondary Experienced Emotion. I had no a priori plans to use responses to the secondary experienced emotion questions in deciding whether to use a picture in Studies 2b and 3. However, I did identify the modal secondary experienced emotion for each picture after I had selected my 12 pictures, to aid with interpreting results in Study 3.

Results

Primary Experienced Emotion

Table 2 shows the modal primary experienced emotion for each picture and the numerical modes.

Table 2

Modal primary experienced emotions for each picture and numerical modes in Study 2a. Bolded rows indicate the pictures that I selected for Studies 2b and 3

Picture Number	Intended Discrete Emotion	Modal Primary Experienced Emotion	Mode of Primary Experienced Emotion
1	Anger	Alert	13
2	**Anger**	**Anger**	**16**
3	Anger	Anger	12
1	**Sadness**	**Sadness**	**42**
2	Sadness	Sadness	42
3	Sadness	Sadness	33
1	Fear	Fear	23
2	**Fear**	**Fear**	**26**
3	Fear	Fear	29
1	Happiness	Happiness	32
2	**Happiness**	**Happiness**	**38**
3	Happiness	Happiness	34
1	Disgust	Disgust	31
2	Disgust	Disgust	45
3	**Disgust**	**Disgust**	**47**
1	**Fatigue**	**Fatigue**	**17**
2	Fatigue	Calm	12
3	Fatigue	Sadness	15
1	**Alert**	**Excited**	**20**
2	Alert	Excited	14
3	Alert	Boredom	9
1	Calm	Calm	23
2	Calm	Calm	22
3	**Calm**	**Calm**	**29**
1	**Excited**	**Excited**	**36**
2	Excited	Excited	27

3	Excited	Happiness	15
1	**Relaxed**	**Relaxed**	**22**
2	Relaxed	Relaxed	20
3	Relaxed	Calm	22
1	Contented	Happiness	25
2	**Contented**	**Happiness**	**22**
3	Contented	Calm	22
1	Bored	I didn't experience any emotion	20
2	**Bored**	**Bored**	**19**
3	Bored	I didn't experience any emotion	20

Below, I detail how I selected the picture for each discrete emotion.

Anger. As seen in Table 2, anger was the primary modal experienced emotion in two of three pictures. I selected picture #2 as the mode for anger was the highest for this picture.

Sadness. Sadness was the primary modal experienced emotion for all three pictures. The modes for sadness were the same for pictures #1 and #2. I selected picture #2 as it: 1) had been validated to induce sadness in the NAPS (Marchewka et al., 2013), and 2) induced the experienced of fewer primary emotions than picture #1. Thus, it was more likely that sadness (as opposed to other emotions) was the primary experienced emotion for most people for this picture.

Fear. Fear was the primary modal experienced emotion for all three pictures. I selected picture #2 rather than picture #3 (whose mode was the highest), as picture #3 depicted a woman being attacked by a man with a knife to her throat, and picture #2 was a picture of a wolf snarling and baring its teeth. The picture of the woman being attacked elicited more anger (N_{anger} = 10) than both the other pictures (Picture #1 N_{anger} = 1; Picture #2 N_{anger} = 1). Thus, I selected the picture of the wolf over the picture of the woman so as to have a picture that induced the experience primarily of fear (as opposed to other emotions).

Happiness. Happiness was the primary modal experienced emotion for all three pictures. I selected picture #2 as the mode for happiness was the highest for this picture.

Disgust. Disgust was the primary modal experienced emotion for all three pictures. I selected picture #3 as the mode for disgust was the highest for this picture.

Fatigue. Fatigue was the primary modal experienced emotion for picture #1 but not the other two pictures. Thus, I selected picture #1 for fatigue.

Alert. Alert was not the primary modal experienced emotion for any of the three pictures. While it was the primary experienced emotion for anger's picture #1, I decided not to use this since it depicted a man with a gun outside a school. According to my conceptualization, alert belonged to the high arousal positive quadrant of the affective circumplex. Thus, I did not select this picture since it had clear negative connotations. I proceeded to select picture #1 since it had the highest number of people who rated alert as the primary emotion ($N_{\text{alert}} = 10$) as compared to picture # 2 ($N_{\text{alert}} = 2$) and picture #3 ($N_{\text{alert}} = 8$), despite not having alert as the primary modal experienced emotion.

Calm. Calm was the primary modal experienced emotion for all three pictures. I selected picture #3 as the mode for calm was the highest for this picture.

Excited. Excited was the primary modal experienced emotion in two of three pictures. I selected picture #1 as the mode for excited was the highest for this picture.

Relaxed. Relaxed was the primary modal experienced emotion in two of three pictures. I selected picture #1 as the mode for relaxed was the highest for this picture.

Contented. Contented was not the primary modal experienced emotion for any of the three pictures. Despite this, I proceeded to select picture #2 since it had the highest number of people who rated contented as the primary emotion ($N_{contented} = 11$) as compared to picture #1 ($N_{contented} = 7$) and picture #3 ($N_{contented} = 5$), despite having happy instead of contented as the primary modal experienced emotion.

Bored. Bored was the primary modal experienced emotion in one of three pictures. Thus, I selected picture #2.

Secondary Experienced Emotion

Overall, for a variety of pictures, people reported experiencing no other secondary emotions in addition to the primary emotions. This was the case for the pictures I selected for sadness, disgust, fatigue, alert, and boredom. For the remaining pictures, people did report experiencing secondary emotions in addition to the primary emotion. In particular, the picture I selected for anger also elicited other secondary emotions (anger, disgust, sadness, and no other emotions). The picture I selected for fear also elicited a feeling of alert. The picture I selected for happiness also elicited excitement. The picture for calm also elicited relaxation and the picture for relaxation elicited calmness. The picture for excitement also elicited happiness. The picture for contentment also elicited calmness.

Overall Reflection

Overall, most participants reported facing no issues throughout the study. Most participants reported that the study was “easy to understand,” “smooth,” that they liked it, and that they experienced “no confusion or difficulty.” I did not need to exclude participants or make any changes to the procedures of the study.

Study 2b

The goal of this study was to validate the picture descriptions to be used in Study 3. Specifically, I conducted this study to ensure that picture descriptions were eliciting the *expectation* of experiencing the specific discrete emotions of interest so I could examine affective forecasting errors in Study 3. Since each of the 12 pictures selected through Study 2a were validated to induce one specific discrete emotion through Study 2a, it was important to have 12 picture descriptions corresponding to each discrete emotion that induced the expectation of that emotion.

Method

Participants

I recruited 50 adult participants (ages 18 years and up) through Prolific. I determined my sample size based on existing financial resources. I only recruited participants who reported U.S. nationality, were residing in the United States, were fluent in English, had a minimum "approval rate" of 95% on Prolific, had not participated in my lab's previous affective forecasting studies, and were using a laptop/desktop (and not a mobile device) to complete the study. Participants were compensated with $5 USD for their participation.

One participant reported that they were not currently in the U.S. and one participant reported that they had previously participated in the study. These participants were excluded. Thus, I had a final dataset of $N = 48$. Please see Table S4 in the Supplementary Materials for the sample's demographic information.

Materials

Pictures. I used the 12 pictures that were validated in Study 2a to induce anger, sadness, fear, happiness, disgust, fatigue, alert, calm, excited, relaxed, contentment, and boredom.

Picture Descriptions. I developed brief descriptions of the 12 selected pictures from Study 2a for use in the Affective Forecasting Task during Part 1 of Study 3. I used ChatGPT (OpenAI, 2023) to generate draft descriptions and then edited these descriptions for accuracy.

Descriptions were 14 to 27 words long ($M = 20.75$, $SD = 4.29$), not including the description stem (i.e., "Imagine a picture of") that was appended to the beginning of every description. Descriptions did not reference the target emotion explicitly. Example descriptions included: "Imagine a picture of a frail, elderly man lying motionless in bed, his hollow cheeks and sunken eyes reflecting a long, painful struggle with illness" (description for sadness) and "Imagine a picture of a crowd of people at a concert waving their hands energetically as colorful confetti fills the air, capturing the thrill of the moment" (description for excitement). All the picture descriptions are provided in the Supplementary Materials.

Self-Report Measures

Primary Expected Emotion. After participants viewed each picture description, they were asked to "select the primary emotion that you expect to experience when viewing the picture" from a list of 14 options including anger, sadness, fear, happiness, disgust, fatigue, alert, calm, excited, relaxed, contented, boredom, other (please specify), and I don't expect to experience any emotion.

Secondary Expected Emotions. After participants answered the question about their primary experienced emotion, they were asked "aside from the primary emotion, what, if any, other emotions do you expect to experience when viewing the picture?" from the same list of 14 response options as before.

Similarity Ratings. After participants viewed each picture, they were asked to "rate how similar the picture was to what you had imagined" on a 7-point scale (1 = *not at all similar*, 7 = *extremely similar*) as in Petagna & Wormwood (2023).

Overall Reflection. At the end of the study, participants answered the same open-ended question about how the study went overall and whether there were any points of confusion or difficulty that affected their participation as in Study 2a.

Procedure

All study procedures were conducted online via Qualtrics. At the onset of the study, participants provided informed consent and answered eligibility questions. Participants were told that they were about to read a series of descriptions of pictures and answer questions about their emotions. They were also asked to picture what was described in their mind's eye as vividly as possible. They then saw each picture description for 30 seconds. After viewing each picture description, they reported the primary and any secondary emotions that they expected to experience when viewing the picture. After these ratings, they were told that they would view the picture that they had just read a description of. They then viewed the picture and rated how similar it was to what they had imagined. This cycle (picture description – expected emotion rating – picture viewing – similarity rating) occurred 12 times. Lastly, participants answered a demographics questionnaire, the overall reflection question, and viewed a debriefing form. See Figure 4 for a schematic of study procedures.

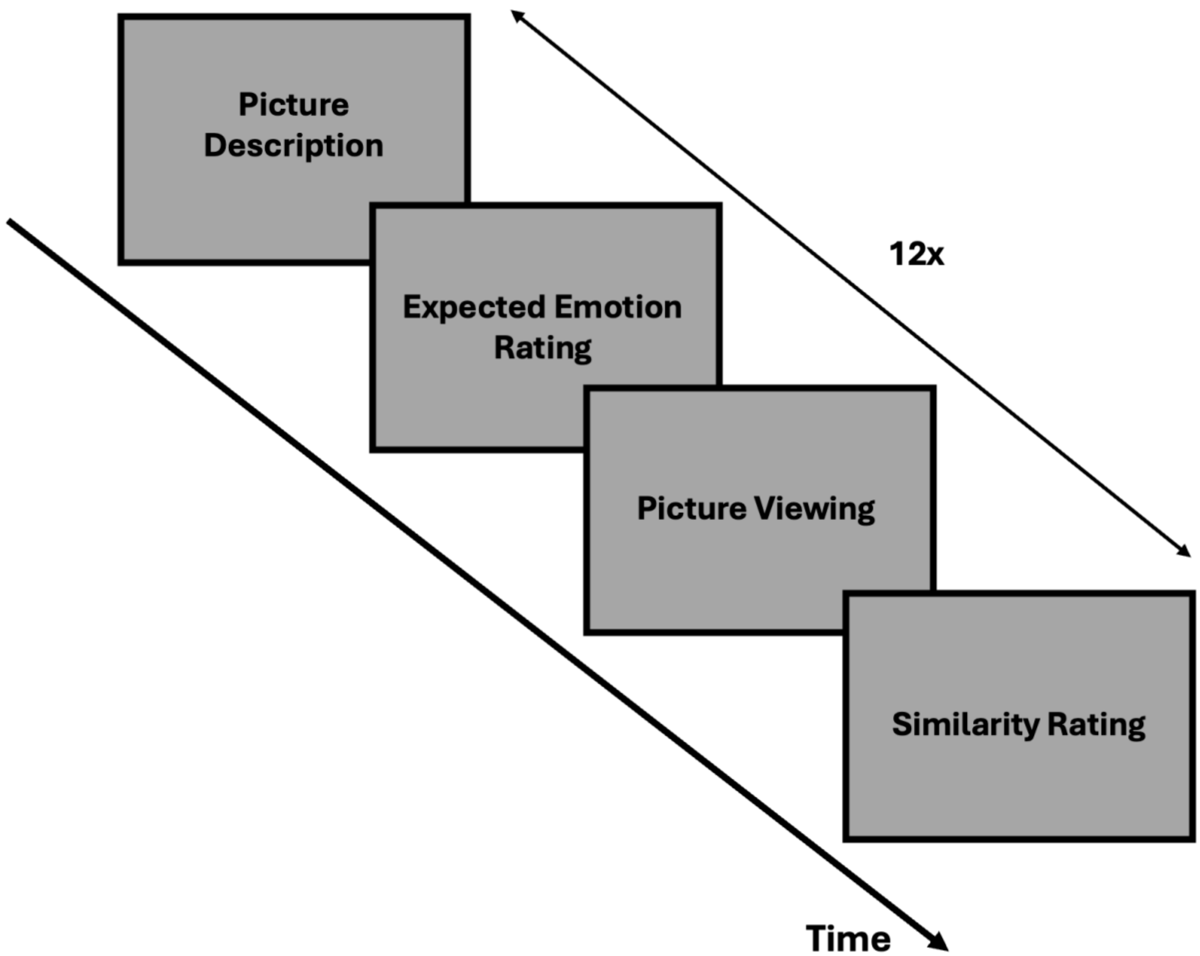

Figure 4. Schematic of Study 2b's procedure

Data Analytic Strategy

Primary Expected Emotion. I identified the modal primary expected emotion for each picture description. If the modal primary expected emotion for each picture description was in fact the expected emotion that I intended to induce (e.g., mode = disgust for the picture description intended to induce the expectation of disgust), I planned to use this picture in Study 3, pending one additional check based on similarity ratings. However, if a picture description was bi-modal (e.g., mode 1 = disgust and mode 2 = sadness for the picture intended to induce the expectation of disgust), I planned to exclude that description from Study 3, or re-run Study 2b with a longer description for that picture.

Secondary Expected Emotion. As in Study 2a, I had no *a priori* plans to use responses to the secondary expected emotion questions in deciding whether to use a picture description in Study 3. However, I did identify the modal secondary expected emotion for each of 12 selected picture descriptions to aid with interpreting results in Study 3.

Similarity Ratings. I conducted individual one sample *t*-tests for the similarity ratings for each of the pictures. If the mean similarity rating for a picture was significantly higher than the lowest possible rating on the scale (i.e., 1), I planned to use this picture description in Study 3. If this was not the case, I planned to exclude this picture description from Study 3.

Results

Primary Expected Emotion

Table 3 presents the modal primary expected emotion for each picture description.

Table 3

Modal primary expected emotions and numerical modes for each picture description in Study 2b

Serial Number	Intended Discrete Emotion	Modal Primary Expected Emotion	Mode of Primary Expected Emotion
1	Anger	Anger	26
2	Sadness	Sadness	42
3	Fear	Fear	30
4	Happiness	Happiness	36
5	Disgust	Disgust	45
6	Fatigue	Fatigue	26
7	Alert	Excited	18
8	Calm	Calm	26
9	Excited	Excited	31
10	Relaxed	Calm	18
11	Contentment	Happiness	16
12	Boredom	Boredom, Calm	16, 16

Overall, as seen in Table 3, eight of the twelve picture descriptions induced the expectation of the intended discrete emotion. This applied to the emotions of anger, sadness, fear, happiness, disgust, fatigue, calm, and excited. Thus, I retained these descriptions for Study 3. The remaining four descriptions for alert, relaxed, contented, and bored also induced the expectation of the intended discrete emotions, albeit with some caveats as detailed below.

Alert. Excited was the primary modal expected emotion for the picture description intended to induce the expectation of alertness. However, the number of participants who selected alert as the primary expected emotion for this description was similar to that of excitement ($N_{alert} = 17$; $N_{excitement} = 18$), which are similar states. Thus, I retained this description for alertness for Study 3, knowing that alert was still a highly relevant emotion for this description.

Relaxed. Calm was the primary modal expected emotion for the picture description intended to induce the expectation of relaxation. However, the number of participants who selected relaxed as the primary expected emotion for this description were approximately the same as those who selected calm ($N_{relaxed} = 17$; $N_{calm} = 18$), which again are similar states. Thus, I retained this description for relaxation for Study 3, knowing that relaxed was still a highly relevant emotion for this description.

Contented. Happiness was the primary modal expected emotion for the picture description intended to induce the expectation of contentment. However, the number of participants who selected contentment as the primary expected emotion for this description was similar to that of happiness ($N_{contentment} = 13$; $N_{happiness} = 16$), which again are similar states. Thus, I retained this description for contentment for Study 3, knowing that contentment was still a relevant emotion for this description.

Bored. Boredom *and* calmness were the primary modal expected emotions for the picture description intended to induce the expectation of boredom. While I had planned to exclude descriptions that were bi-modal, states that were overlapping were hard to isolate and measure. Furthermore, I was unable to repeat this study with one new, richer, and longer picture description for boredom as 1) increasing the length of one description would make it systematically longer than the other 11 descriptions and might introduce a confound into Study 3, and 2) this was logistically not feasible given my existing resources and timeline. Thus, given that a third of the sample reported that they expected to experience primarily boredom, and another third of the sample reported that they expected to experience primarily calmness, I retained this description for boredom for Study 3, knowing that boredom was still a highly relevant emotion for this description.

Secondary Expected Emotion

For some picture descriptions, people reported expecting to experience no secondary emotions in addition to the primary emotions. This was the case for the picture descriptions for sadness, fatigue, and alertness. For the remaining picture descriptions, people did report expecting to experience secondary emotions. In particular, the picture description I selected to induce the expectation of anger also elicited the expectation of disgust. The picture description I selected to induce the expectation of fear also elicited the expectation of alertness. The picture description I selected to induce the expectation of happiness also elicited the expectation of contentment and the description for contentment also induced the expectation of happiness. The picture description I selected to induce the expectation of disgust also elicited the expectation of anger. The picture description I selected to induce the expectation of calmness also elicited the expectation of contentment. The picture description I selected to induce the expectation of excitement also elicited the expectation of happiness. The picture description I selected to induce the expectation of boredom also elicited the expectation of calmness.

Similarity Ratings

As seen in Table 4, all the picture descriptions passed the similarity check. That is, I rejected the null hypothesis that the picture that participants had imagined in their mind based on each of the picture descriptions was not at all similar to the picture that they viewed (i.e., equal to the lowest rating on the scale). In fact, these mental pictures were rather similar to the actual pictures as seen through the mean similarity ratings which ranged from 4.50 - 6.06 on a 7-point scale with 7 being “extremely similar.”

Table 4

Results from similarity rating analyses for each picture description in Study 2b

Picture Description	*t*	*df*	*M*	*95% CI*	*p*
Anger	16.85	47	4.88	4.41, 5.33	<.001***
Sadness	14.66	47	4.85	4.33, 5.38	<.001***
Fear	19.13	47	5.52	5.05, 6.00	<.001***
Happiness	27.34	47	5.73	5.38, 6.08	<.001***
Disgust	24.95	47	5.83	5.44, 6.22	<.001***
Fatigue	23.39	47	5.65	5.25, 6.04	<.001***
Alert	30.34	47	6.06	5.73, 6.40	<.001***
Calm	23.94	47	5.65	5.26, 6.04	<.001***
Excited	26.20	47	5.73	5.37, 6.09	<.001***
Relaxed	24.70	47	5.58	5.21, 5.96	<.001***
Contentment	21.91	47	5.56	5.14, 5.98	<.001***
Boredom	14.26	47	4.50	4.01, 4.99	<.001***

Note. *$p < .05$; **$p < .01$; ***$p < .001$

Overall Reflection

As in Study 2a, most participants reported facing no issues throughout the study. Most participants reported that the study was “great,” “very fun,” that it “went well with no problems” and that it was “interesting.” One participant reported that the photos were similar to what they had envisioned. Another participant reported that they would occasionally picture themselves in the images that they created in their head and then re-oriented their train of thought back to the description. I did not need to exclude participants or make any changes to the procedures of the study.

Studies 2a and 2b: Discussion

The goal of Studies 2a and 2b was to validate the pictures and picture descriptions that would be used in Study 3 to induce the experience (Study 2a) and expectation (Study 2b) of individual discrete emotions. This was in contrast to Study 1, where Petagna and Wormwood selected pictures from the IAPS that were validated to induce four dimensional states.

One of the drawbacks of using Petagna and Wormwood's (2023) data for the purpose of Study 1 was that their stimuli norming approach was dimensional; however, they measured several discrete emotions, some of which may not have been applicable to or very highly rated for each picture. This could have ultimately prevented us from detecting differences based on measurement approach. Thus, in order to address this limitation in Study 3, I conducted Studies 2a and 2b to validate our stimuli to induce *one primary discrete emotion*.

Through Study 2a, I was confident that the pictures for ten of the selected discrete emotions (anger, sadness, fear, happiness, disgust, fatigue, calm, excited, relaxed, bored) would induce the experience of the intended emotions. I was less confident that pictures for the remaining two emotions (alert, contented) would induce the experience of the intended emotions. However, I chose to retain them in the stimuli set and planned to repeat any confirmatory analyses with and without these pictures and report whether conclusions changed.

Through Study 2b, I was confident that the picture descriptions for eight of the selected discrete emotions (anger, sadness, fear, happiness, disgust, fatigue, calm, excited) would induce the expectation of the intended discrete emotion. I was less confident that picture descriptions for the remaining four emotions (alert, relaxed, contented, bored) would induce the expectation of the intended emotions for *everyone* in the sample. However, I chose to retain them in the stimuli set, repeat any confirmatory analyses without these picture descriptions, and report whether conclusions changed.

In sum, I concluded that most of the 12 pictures and picture descriptions would induce the experience or expectation of *one intended primary* emotion or dimensional state. While responses to the questions about secondary emotions suggested that some pictures and picture descriptions also induced the experience and expectation of other emotions, this was to be expected. As long as most participants reported experiencing or expecting to experience the intended *primary* emotion, I concluded that the construct validity for a single item question in Study 3 that measured forecasts and actual affect ratings about this primary emotion was not threatened. Thus, with this newly-validated set of pictures and picture descriptions, I was equipped to detect any differences between affective forecasting errors as measured by discrete and dimensional approaches in Study 3.

Study 3

The goal of Study 3 was to build on and extend the findings of Study 1 while also addressing its limitations. Specifically, participants were randomly assigned to a discrete or dimensional condition, viewed descriptions of pictures, made forecasts about discrete emotions or how positive or negative they expected to feel, viewed the pictures that matched the descriptions, and rated how they actually felt.

Study 3 built on Study 1 in several ways. First, Study 3 employed a between-subjects design and ensured that the processes underlying ratings of affect using discrete and dimensional measures were not conflated. Second, Study 3 recruited a sample of adults across the U.S., thus improving generalizability. Third, this study used stimuli validated in Studies 2a and 2b to elicit individual discrete emotions. Fourth, unlike in Study 1, this study measured how negative/positive (Watson and Tellegen, 1985) people expected to feel/actually felt. I selected this measure over measuring valence and arousal (Petagna and Wormwood, 2023) for two reasons: 1) to use the dimensional approach in a new way, given the null results from Study 1's use of the dimensional approach, and 2) to parallel the measure of discrete emotions as closely as possible without introducing a confound. Since I asked participants to make ratings about *one* discrete emotion (e.g., disgust) per picture in the discrete condition, asking participants to make ratings about *two* dimensional states (e.g., valence and arousal) would have introduced a methodological confound.

Given the gaps in the existing literature and limitations of Study 1, the primary research question was as follows: Do affective forecasting errors for discrete emotions versus dimensional affective states differ? To answer this question, I used a between-subjects design to measure forecasting errors for discrete and dimensional states. Additionally, as in Study 1, I included measurement type and type of affect rating as factors in my model and examined whether directional affective forecasting errors differed based on measurement type (discrete/dimensional). Furthermore, as in Study 1, I explored whether absolute affective forecasting errors differed based on measurement type, and whether the relationship between affective forecasting errors and measurement type was moderated by a series of other variables.

If I found that affective forecasting errors were different for discrete emotions than for dimensional affective states, this would suggest that some of the variation in existing findings might be accounted for by measurement approaches. Additionally, this would indicate that the processes involved in predicting and/or generating discrete emotions are somehow different than the processes involved in predicting and/or generating dimensional affective states. Future research would be needed to understand those differences. In contrast, if I found, as in Study 1, that affective forecasting errors did not differ for discrete emotions than for affective dimensional affective states, this would indicate that something other than variation in measurement approaches explains variation in affective forecasting errors across studies. Future research would be needed to identify other sources of variation. Alternatively, null findings could also suggest that aspects of the study design were obscuring any true differences between affective forecasting errors for discrete versus dimensional states. I did not have *a priori* predictions about which approach would show larger errors than the other, or the direction of those errors.

Method

Design

This study used a between-subjects design. There were two predictors of interest. The predictor variables included type of affect rating (predicted, actual) and measurement type (discrete, dimensional). Type of affect rating was manipulated within subjects. Measurement type was manipulated between subjects. Participants were randomly assigned to a discrete or dimensional condition. The dependent variable was the level of affect rating.

Participants

Given the timeline of this project as well as existing financial resources, I aimed to recruit approximately 200 total participants (n = 100 per measurement type). Based on an *a priori* power analysis, 135 participants would afford 80% power to detect an interaction between type of affect rating and measurement type with at least a small effect size (d = 0.25) (Cohen, 1988). Thus, a sample of N = 200 was more than adequate to detect small effects.

I ultimately recruited 205 adult participants (ages 18 years and up) through Prolific.[3] I only recruited participants who reported U.S. nationality, were residing in the United States, were fluent in English, had a minimum "approval rate" of 95% on Prolific, had not participated in my lab's previous affective forecasting studies, and were using a laptop/desktop (and not a mobile device) to complete the study. Participants were compensated with $5 USD for their participation in the study.

[3] The sample I ultimately recruited (N = 205) deviated slightly from the sample I planned to recruit in the preregistration (N = 200). Through a Graduate Student Research Competition grant from Tufts University, I had the financial resources to recruit 5 additional participants, which could prove useful assuming some data would likely be excluded.

Four participants reported that they had previously participated in the study. These participants were excluded. Thus, I had a final dataset of *N* = 201. Please see Table S6 in the Supplementary Materials for the sample's demographic information. All study procedures were approved by the Social, Behavioral, and Educational Institutional Review Board at Tufts University prior to data collection.

Self-report Measures

Baseline affect. At the start of the study, all participants were asked to "describe how <adjective> you are CURRENTLY feeling" on a 7-point scale (1 = *not at all*, 7 = *extremely*). Specifically, they rated all 12 discrete emotions (anger, sadness, fear, happiness, disgust, fatigue, alert, calm, excited, relaxed, contented, boredom), and two dimensional states (positive, negative).

I subtracted the sum of the ratings of the negative emotions and affective states from the sum of the positive emotions and affective states. Positive baseline affect scores indicated that participants experienced greater positive than negative emotions/affect at baseline. Negative scores indicated that participants experienced greater negative than positive emotions/affect at baseline.

Discrete emotions. Approximately half the participants rated the 12 discrete emotions that corresponded to each description of a picture and an actual picture. Participants were asked to "predict how <adjective> you think you will feel when viewing the picture" on a 7-point scale (1 = *not at all*, 7 = *extremely*). They were also asked to "rate how <adjective> you felt when viewing the picture" on the same 7-point scale.

Dimensional states. The other half of participants rated how positive or how negative they felt when exposed to positive or negative, respectively, descriptions and pictures. Participants were asked to "predict how positive/negative you think you will feel when viewing the picture" on a 7-point scale (1 = *not at all*, 7 = *extremely*). They were also asked to "rate how positive/negative you felt when viewing the picture" on the same 7-point scale.

Overall Reflection.[4] At the end of the study, participants answered the same open-ended question about how the study went overall and whether there were any points of confusion or difficulty that affected their participation as in Studies 2a and 2b.

Distraction Task

Participants completed a relatively easy Sudoku task for five minutes as a distraction. In this task, participants entered numbers into blank cells within a grid that was pre-populated with some numbers in certain areas of the grid. The purpose of this task was to ensure that participants did not remember the affective forecasts they had just made. The Sudoku task had the benefit of using numbers that anchor the affective ratings scale, which may have further prevented participants from remembering their affective forecasts. Performance data from this task was not collected. Data about how long it took participants to click during the task for the first time, last time, and to click the submit button was collected. Additionally, data about the number of clicks they made during the task was collected. I included the number of clicks as a covariate in my exploratory analysis.

Attention Check

Participants completed two attention checks. Specifically, I included an extra neutral description of a picture and actual picture (the picture of a ceiling fan) at the end of each of the phases. Participants were asked to "select the option that best describes the picture description that you just viewed (Phase 1) or the picture that you just viewed (Phase 2)." They saw four

[4] Note that the procedure I preregistered for this study did not include this Overall Reflection question. However, after the insights that were collected in Studies 2a and 2b, I decided to deviate from the preregistered procedure and include this question in Study 3 as well.

options, one of which accurately described the picture; the remaining options were unrelated to the picture. I excluded data from participants who did not select the correct answer in *both* attention checks ($N = 9$) in our confirmatory analyses.

Stimuli

Pictures. I used the 12 pictures identified through Study 2a above to induce the selected high and low arousal positive and negative discrete emotions.

Picture Descriptions. I used the 12 picture descriptions identified through Study 2b above to induce expectations of the selected high and low arousal positive and negative discrete emotions.

Procedure

All study procedures were conducted online via Qualtrics. At the onset of the study, participants provided informed consent and answered eligibility questions. Next, all participants reported their baseline affect. After this, participants were randomly assigned to one of two conditions (discrete, dimensional) through Qualtrics with the constraint to present conditions evenly. Procedures then unfolded in two phases.

In Phase 1, participants completed an Affective Forecasting Task (as in Petagna & Wormwood, 2023). Specifically, they were told that they were about to read a series of descriptions of pictures and answer questions about their emotions. They were also asked to picture what was described in their mind's eye as vividly as possible. In this task, participants viewed 12 picture descriptions in random order. Each description was presented for 30 seconds.

After each description, participants were asked to predict how they thought they would feel when viewing the picture described on a 7-point scale (1 = *not at all*, 7 = *extremely*). Depending on the condition to which they were assigned, participants predicted one discrete emotion (e.g., sadness if sadness was the target emotion) or one matched dimensional affective state (e.g., negative if the sadness was the target emotion). Participants were able to advance once they made their predictions (untimed). At the end of these 12 descriptions and ratings, they completed the first attention check. After this, participants completed a Sudoku task for five minutes. This concluded Phase 1 of the Study.

In Phase 2, participants complete a Picture Viewing and Affect Rating Task (as in Petagna & Wormwood, 2023). Specifically, they were told that they would view the previously-described pictures and answer questions about their emotions. In this task, participants viewed the 12 pictures that were previously described in random order. Each picture was displayed for 12 seconds. Participants rated how they actually felt when viewing each picture using the same discrete or dimensional measures as in Phase 1. Participants were able to advance once they had made their ratings (untimed). At the end of these 12 pictures and ratings, they completed the second attention check. This concluded Phase 2.

Lastly, participants completed a demographics questionnaire, the overall reflection question, and viewed a debriefing form. See Figure 5 for a schematic of the main study procedures.

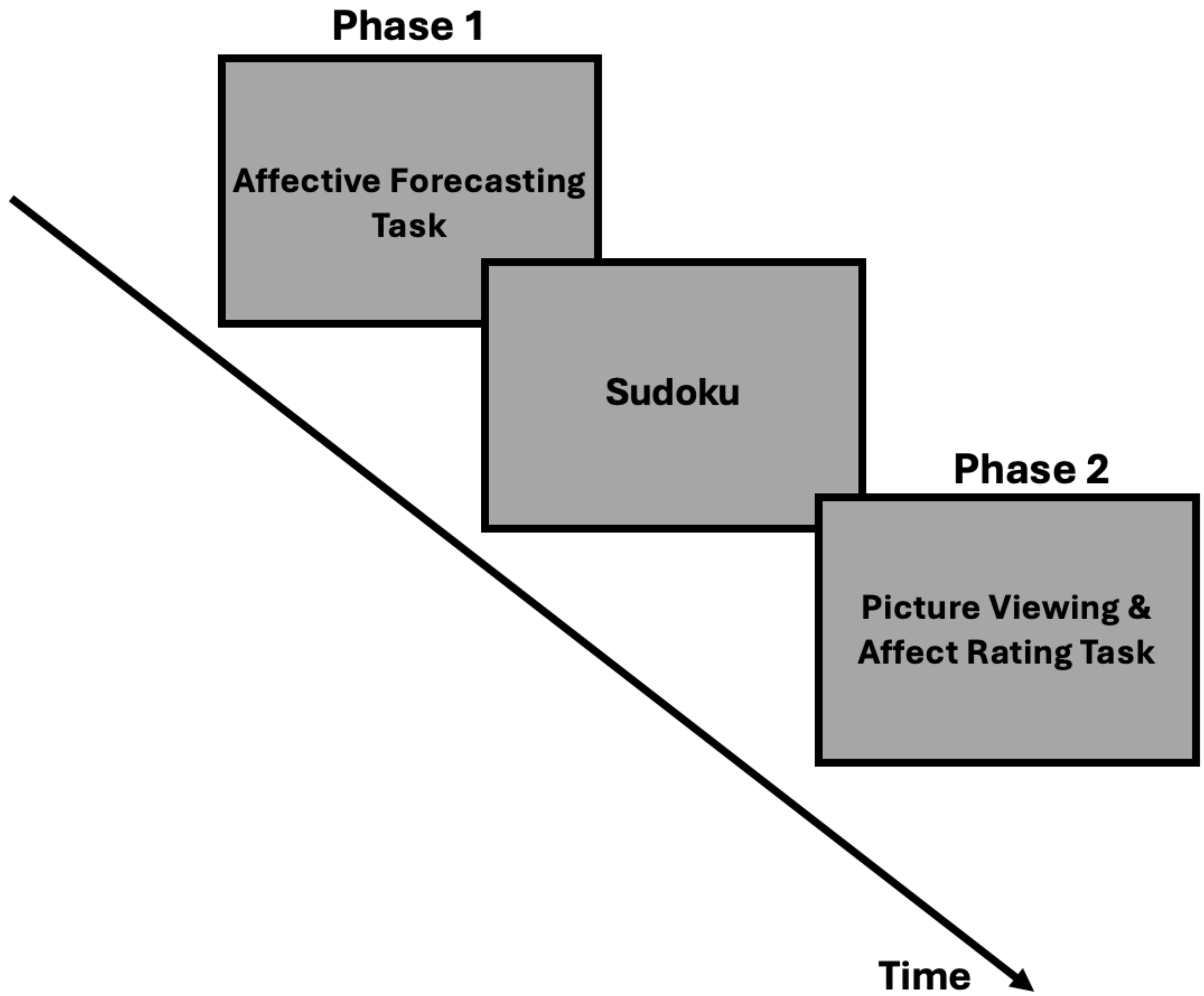

Figure 5. Schematic of Study 3's main procedures

Data Analytic Strategy

To test whether discrete and dimensional measurement approaches predicted unique patterns of affective forecasting errors, I conducted a linear mixed effects regression examining sources of variation in ratings. I used *R* (Version 4.4.1) to conduct this model with the *lme4* package (Bates et al., 2015).

The model was structured similarly to the model for Study 1. It included fixed effects for type of affect ratings, measurement type, and their interaction, as well as random intercepts for participants and items (picture description + picture combination). Including the random intercepts for participants and items allowed the model to detect variation in the data that could be accounted for by participants making different ratings and items on the scale being different from one another. I pre-registered a model that also included a by-participant random slope for type of affect rating to examine whether different participants exhibited different levels of affective forecasting errors. While the model converged with the by-participant random slope and random intercept for participants, it resulted in singular fit because the correlation between the slope and intercept was 1. This suggested that the random slope was redundant and arguably explaining the same variation as the random intercept. Thus, I deviated from the pre-registered data analysis plan and removed the by-participants random slope from the confirmatory model. Additionally, I deviation-coded type of affect ratings (predicted affect ratings = 0.5, actual affect ratings = -0.5) and measurement type (dimensional = 0.5, discrete = -0.5).

I first ran an empty model without the fixed effects to examine the percentage of the total variance in affect ratings that was explained by participants and items. I found that 23.6% of the variation was explained by differences among participants in the sample. Additionally, 28.1% of the variation in ratings was explained by items, suggesting that some of the variation in ratings was accounted for by the items being different from one another. This suggested that there was clustering in the data and that utilizing multilevel modeling was an appropriate analysis method. Therefore, I retained the random intercepts for participants and items in my confirmatory model.[5]

I expected that the magnitude of affective forecasting errors would differ based on whether they were measured using a discrete versus a dimensional approach. However, I did not have specific *a priori* predictions about which approach would show larger errors than the other,

---

[5] The analysis was conducted using the *lme4* package (Bates et al., 2015) in R and utilized the following structure:

```
lmer(rating ~ affect_type*measure_type + (1 |ID) + (1|item),
        dataframe,
        contrasts = list(affect_type = affect_type.deviation.con,
                         measure_type = measure_type.deviation.con))
```

or the direction of errors. A two-way interaction between type of rating and measurement approach would be required to conclude support for the hypothesis. Lastly, if the hypothesized interaction between type of rating and measurement approach from the linear mixed effect regression was not statistically significant, I planned to examine whether it was equivalent to zero with equivalence bounds of -0.10 to 0.10 in standardized units.[6]

Results

Confirmatory Analysis

Overall, as seen in Table 5, the fixed-effect predictor variables (type of affect rating, measurement type, and their interaction) collectively accounted for 0.2% of the variance in affect ratings, the dependent variable.

---

[6] I used the equivalence_test() function in the parameters package in R as follows:

```
equivalence_test(
  confirmatory_model,
  range = "default",
  ci = 0.95,
  rule = "classic",
  effects = c("fixed", "random"),
  verbose = TRUE)
```

Table 5

Results from linear mixed effects regression on affect ratings, including fixed effects for type of affect rating, measurement type (both deviation-coded), and random intercepts for participants and items

	Affect Rating		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	4.82	4.22 – 5.42	<0.001***
Predicted-Actual	0.10	0.03 – 0.18	0.009**
Dimensional-Discrete	-0.12	-0.40 – 0.16	0.416
Predicted-Actual × Dimensional-Discrete	-0.02	-0.17 – 0.14	0.829
Random Effects			
σ^2		1.84	
τ_{00} PROLIFIC_PID		0.90	
τ_{00} items		1.07	
ICC		0.52	
N PROLIFIC_PID		192	
N items		12	
Observations		4607	
Marginal R^2 / Conditional R^2		0.002 / 0.519	

Note. *p < .05; **p <.01; ***p < .001

There was a significant fixed effect of type of affect rating, suggesting that participants predicted (M = 4.90, *95% CI* [4.30, 5.50) significantly greater emotions or affect than they actually experienced (M = 4.80, *95% CI* [4.20, 5.40]), B = 0.11, z = 2.62, p = .009, regardless of measurement type.

However, results from the equivalence test indicated that the fixed effect of type of affect rating was small enough to be equivalent to zero, with a 95% CI for the effect size falling within the ROPE [-0.19, 0.19], *95% CI* [0.04, 0.17], and a statistically significant *p*-value of .014. The SGPV was 0.996, indicating strong evidence for equivalence. This suggests that although the difference between predicted and actual emotion/affect deviated significantly from 0, it was nevertheless a very small effect.

The hypothesized two-way interaction between type of affect rating and measurement type was not statistically significant (see Figure 6).[7]

Results of the equivalence test indicated that the interaction between type of affect rating and measurement type was small enough to be equivalent to zero, with a 95% CI for the effect size falling within the ROPE [-0.19, 0.19], *95% CI* [-0.15, 0.11], and a statistically significant *p*-value of .018. The SGPV was 0.995, indicating strong evidence for equivalence. Contrary to the confirmatory hypothesis, participants in the discrete and dimensional conditions exhibited similar affective forecasting errors.

[7] After applying pre-registered exclusions, one participant in the sample reported having color-blindness. In case this participant's responses unduly influenced our results and conclusions, I repeated the confirmatory analysis without this participant. Results and conclusions remained unchanged; thus, the results reported in the main text include the entire sample.

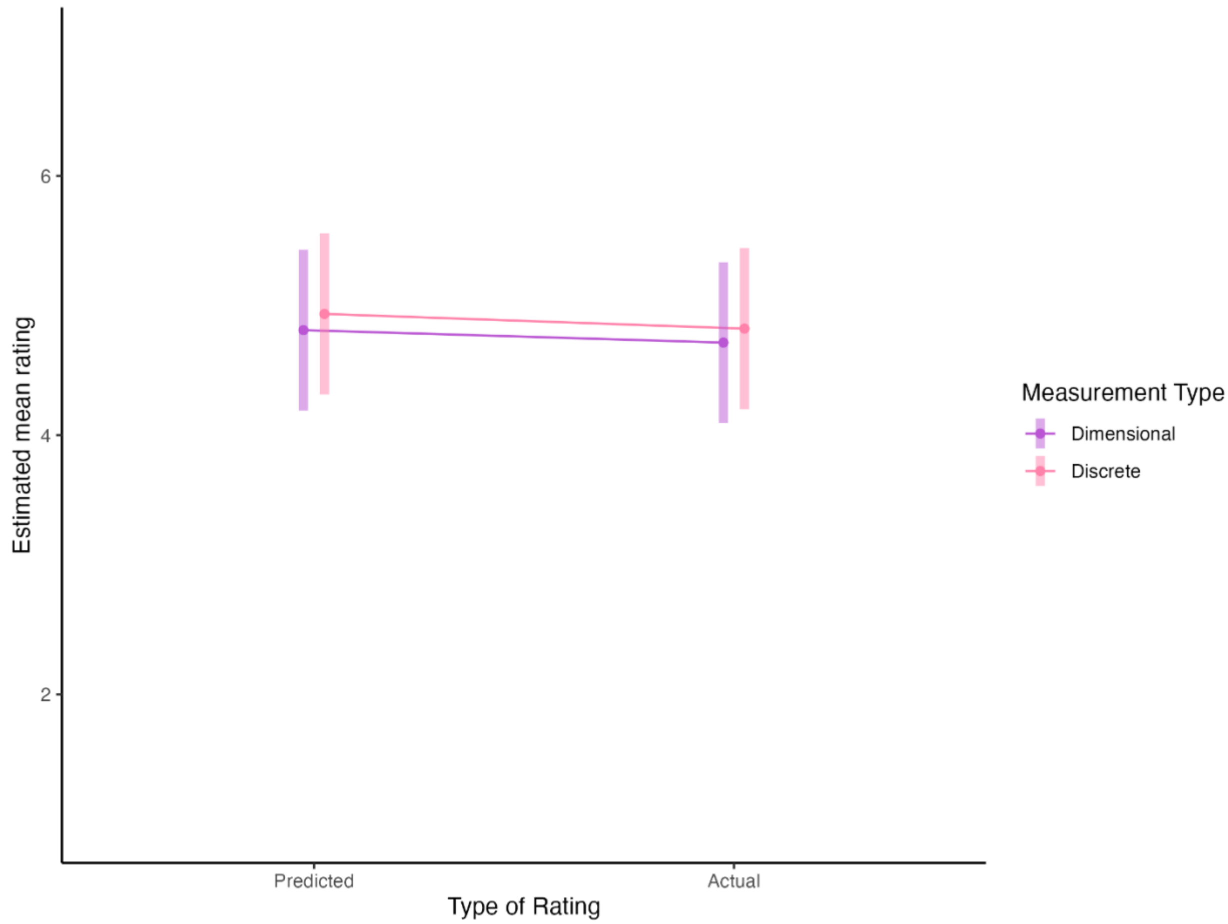

Figure 6. Estimated mean affect ratings as a function of measurement type conditions (dimensional, discrete) and type of affect rating (predicted, actual). Error bars represent 95% confidence intervals around each mean, shown in filled circles.

Overall Reflection

Most participants reported facing no issues throughout the study. A majority reported that the study "went smoothly," that they enjoyed it, that it was an "interesting experience," that "instructions were clear," and that the topics were "relevant and engaging" and allowed for "thoughtful responses." A handful of participants also mentioned the Sudoku task in their responses. For example, some participants reported enjoying this task, while others reported never having tried it previously, and yet others reported that they were not very good at the task. I did not need to exclude participants or make any changes to the procedures of the study.

Exploratory Analysis

The goal of Study 3 was to examine *directional* affective forecasting errors – under- or overestimation – through raw affect rating scores. However, given that Petagna & Wormwood (2023) and other affective forecasting researchers (e.g., Frank et al., 2021) have often focused on the *magnitude* of errors by calculating the absolute value of affective forecasting error scores, I also applied the confirmatory model on absolute affective forecasting scores (i.e., |predicted-actual affect|). Consistent with the confirmatory analysis, I found no statistically significant effect of measurement type on absolute affective forecasting error scores. Overall absolute affective forecasting errors were about 1 unit ($B_0 = 0.82$, *95% CI* [0.72, 0.92], $p < .001$), and the distribution was positively skewed. Please see Table S7 in the Supplementary Materials for the results of this model. Lastly, unlike the confirmatory analysis on directional affective forecasting errors, results from the equivalence test indicated that the absolute affective forecasting errors were not equivalent to zero, with a 95% CI for the effect size falling outside the ROPE [-0.10, 0.10], *95% CI* [0.73, 0.91], and a non-significant *p*-value of >.999. The SGPV was <.001, indicating weak evidence for equivalence.

Furthermore, I explored a series of additional models including the following predictors - baseline affect (centered), age (centered), number of clicks during the Sudoku task, stimulus valence (pleasant, unpleasant), and stimulus arousal (high, low) - as covariates in the confirmatory model. I did so to examine whether these variables moderated the interaction between type of affect rating and measurement type. This would allow them to soak up part of the variation that was previously being considered noise, thereby increasing sensitivity to detect any interactive effects of type of affect rating and measurement type. Overall, I did not find any statistically significant interactions between any of these covariates and type of affect rating and measurement type. Please see Tables S8 through S11 in the Supplementary Materials for the results of these models.

According to my pre-registered data analysis plan, I also explored a model that *excluded* the pictures and picture descriptions I was not as confident about through Studies 2a and 2b. In particular, I was confident that the pictures and picture descriptions would induce the intended expectation and experience of anger, sadness, fear, happiness, disgust, fatigue, calmness, and excitement. I was less confident about the pictures intended to induce alertness and contentment[8] and the descriptions intended to induce alertness, contentment, boredom, and relaxation. Consistent with the confirmatory analysis, there was no statistically significant interaction between type of affect rating and measurement type on affect ratings when excluding these pictures and descriptions from the confirmatory model. Please see Table S12 in the Supplementary Materials for the results of this model.

Discussion

The goal of Study 3 was to build on and extend the findings of Study 1, address its limitations, and examine whether participants exhibited different levels of affective forecasting errors for discrete and dimensional states using a new approach. Specifically, participants were randomly assigned to a discrete or dimensional condition, viewed descriptions of pictures, made forecasts about discrete emotions or dimensional states, viewed the pictures, and rated how they actually felt. Contrary to the hypothesis, affective forecasting errors did not differ for discrete

[8] I was confident that the picture and picture description for fatigue would induce fatigue. However, I preregistered a plan in which I would exclude responses to the pictures and descriptions for fatigue by mistake. Thus, I deviated from the preregistered plan and included responses to pictures and picture descriptions intended to induce fatigue in this exploratory analysis. Additionally, consistent with my preregistered analysis plan, I excluded responses to the pictures and descriptions for boredom because I was not confident that the picture description would induce boredom.

emotions versus dimensional affective states. Unlike in Study 1, however, I found that participants overestimated the intensity of their response to the pictures, a *directional* affective forecasting error. Participants also exhibited affective forecasting errors in *absolute* terms.

Similar to Study 1, in this Study, I found that there was insufficient variation in affective forecasting errors. This was true in both the raw and absolute data. Insufficient variation was seen in the fact that most values for the predicted - actual differences were near zero. Furthermore, although raw forecasting errors were statistically significant, equivalence test results suggested that these errors were also small enough to also be equivalent to 0. Thus, due to the restricted range of forecasting errors in this study as well, it is possible that we were unable to detect any effects of measurement type, had they existed.

Alternatively, the null results from this study could suggest that the hypothesis is incorrect and that in reality, people are not differentially accurate at predicting discrete versus dimensional states. Furthermore, this could also suggest that a factor other than measurement type may explain differences in forecasting error across studies in the literature, should they exist. However, it is also possible that these null results are due, in part, to some of the limitations of this study as detailed below.

First, participants in this study made both affective forecasts and actual affect ratings in a *single session* within a relatively short time frame (~30 minutes). Unlike in Study 1, where participants made affective forecasts and *one week later* made ratings of actual affect, in Study 3, participants had approximately five minutes between their affective forecasts and actual affect ratings. While participants completed a Sudoku task to distract them and reduce the possibility that they remembered the affective forecasts that they just made, it is still possible that they recalled their forecasts while they rated their actual affect. Participants might subsequently have been influenced to report their actual affect such that it aligned with their affective forecasts.

In fact, some responses from the overall reflections about the study suggested that some people recalled their affective forecasts at the time that they rated their actual affect. For example, one participant reported "… I was surprised by how most of my predictions were correct, but some were wrong once I saw the actual picture…" While this is anecdotal evidence from just one participant, it does support the idea that the interval between affective forecasts and ratings of actual affect might not have been sufficient to make participants forget their forecasts. This might, in turn, explain the small affective forecasting errors I found and also why I may not have seen any effect of measurement type on forecasting errors.

Past studies that have been able to detect affective forecasting errors within a single study session (e.g., Patel & Urry, under review; Coundris et al., 2023) and have found that a one-minute delay is sufficient to make participants forget 30 affective forecasts (Floerke & Urry, unpublished). Given this evidence, limited resources, and time constraints, I did not use a week-long procedure as in Study 1, as I believed that a five-minute delay between 12 affective forecasts and ratings of actual affect would be sufficient. However, it would benefit future researchers to repeat this study with a longer delay between forecasts and ratings of actual affect, to prevent participants, such as the one referenced above, from remembering their forecasts at the time of their actual affect ratings. These researchers might be better equipped to detect an effect of measurement type on affective forecasting errors, should it exist.

Second, while a strength of this study was that it examined affective forecasts for a variety of low and high arousal positive and negative emotions, these emotions and affective states were just a subset of the possible emotions and affective states that people experienced. Specifically, I only examined 12 discrete emotions and two dimensional states in this study. I did so to ensure that I was capturing a large range of emotions that participants were likely to experience while also representing three emotions each from all four quadrants of the affective circumplex. However, it is possible that affective forecasting errors differ by measurement type for discrete emotions that I did not measure in this study, and for more granular dimensional states. For example, had I measured other discrete emotions such as nervousness or other fine-grained dimensional states such as high arousal positive affect, high arousal negative affect, low arousal positive affect, or low arousal negative affect, I might have been better positioned to uncover any differences in forecasting errors that existed. Due to time constraints, I could not examine all the possible discrete emotions and granular dimensional affective states; however, future research may benefit from building on this study by examining other emotions/affective states that I did not. It seems unlikely that researchers will uncover measurement type effects on forecasting errors for other emotions/affective states because I carefully selected the emotions in this study such that they captured all the likely primary emotions. However, this might still be an avenue worth exploring.

Relatedly, while I conducted Studies 2a and 2b to ensure that pictures and descriptions elicited just *one* primary discrete emotion and its corresponding dimensional state, it is likely that participants in Study 3 ultimately experienced more than the one primary discrete emotion/dimensional state. In fact, examining the secondary experienced and expected emotions in Studies 2a and 2b suggested that some participants *did* experience emotions other than the primary intended emotions for some pictures and picture descriptions. Additionally, inspection of responses to the overall reflection question also suggested that despite validating stimuli through Studies 2a and 2b, there was some variation in the emotions that people experienced/expected to experience, and not everyone necessarily experienced the intended emotion. For example, one participant reported this in regard to the picture that was intended to induce fear: “I’m not sure if my response is typical or not, but the wolf one I didn't feel negatively about because I like animals and also for some reason that particular expression did look a little funny on the wolf".

Thus, it is possible, for example, that upon viewing the picture that had been validated to induce fear/negative affect, participants felt fear for the most part (primary emotion), but also slightly happy (secondary emotion). Alternatively, it's possible that *most* participants felt primarily fear, but a handful of them, such as the one above, felt primarily happy. It is possible that measurement type would have affected affective forecasting errors for these secondary emotions/affective states, which I did not capture in the current study. I only focused on forecasting errors for the primary discrete emotion/affective state per picture. I limited the task to one item for each picture for the sake of time, to minimize participant fatigue, and only measure *relevant* emotions instead of a list of 16 emotions as in Study 1 (Petagna and Wormwood, 2023). However, this design choice came with trade-offs. In the future, it could be helpful to employ a mixed-methods approach and ask participants to *identify* the emotions that they expect to experience and actually experience themselves. One could subsequently ask participants to make a rating about *how much* of these emotions they expect to experience and actually experience on a Likert scale. This would ensure an idiographic approach that is tailored to *each* participant and would capture affective forecasting errors for the primary emotion that *each* participant experienced.

Third, in this study, the picture descriptions were longer (M = 20.75, SD = 4.29) than those in Study 1 (M = 12.40, SD = 3.23). I intentionally crafted longer descriptions as I wanted them to be evocative and prevent substantial misconstrual (Wilson & Gilbert, 2003). Misconstrual is one of the sources of affective forecasting errors. It occurs when people imagine a certain event while making an affective forecast, but ultimately end up experiencing an event that does not align with their imagined event. This unexpected nature of the actual event leads to the discrepancy between people's affective forecasts and their actual affect. In creating longer descriptions than Study 1, my goal in this study was to minimize forecasting errors due to misconstrual and create a context in which I could examine forecasting errors due to the measures used to test them. However, it is possible that my efforts to minimize misconstrual *greatly reduced* misconstrual for participants. That is, it might be the case that the picture descriptions were so descriptive that they resulted in accurate depictions in people's mind during affective forecasting, and thus prevented me from finding large affective forecasting errors.

In fact, responses to the similarity question in Study 2b suggested that pictures were quite similar to what participants had imagined after reading picture descriptions. Furthermore, responses to the overall reflection question suggested that for some people, picture descriptions evoked an accurate representation of the pictures they saw. For example, one participant reported "That toilet bowl is just as nasty in the picture as described!" However, another participant reported that "… I did notice that the images I generated myself in my head during the first task turned out to be a lot more evocative than the actual presented images in the third part…". Thus, it is difficult to ascertain whether the picture descriptions erased misconstrual, to what extent it did so, and for how many participants this was the case, since participants' reflections were anecdotal, not coded for systematically, and participants were not explicitly asked about misconstrual or similarity in this study. Nevertheless, to the extent that I might have reduced the contribution of misconstrual to affective forecasting errors, this could have contributed to restriction of range discussed above and null effects of measurement type.

Lastly, while Study 3's between-subjects study had one clear benefit - namely, that participants in each of the conditions only made one type of rating (discrete or dimensional), thus not conflating any of the underlying processes that occurred when they predicted or rated their actual state using discrete and dimensional approaches - this design choice also came with one major drawback. Between-subjects designs do not eliminate individual differences among participants that within-subjects designs tend to minimize. While some of this variation between participants was addressed through random assignment to conditions and the random intercept for participants, this design choice nonetheless added noise to the data and could have made it difficult to detect small differences, were there any. I ultimately decided against a within-subjects approach as I believed that asking participants to make ratings using both discrete and dimensional measures in a relatively short period of time (~1 hour) could conflate and muddy the underlying processes that occurred when people predicted or rated their actual emotions. Furthermore, a within-subjects design would also necessitate a longer study and might have resulted in participants feeling fatigued and not providing accurate responses. In order to fully isolate the underlying processes and test most accurately affective forecasting errors for discrete and dimensional states, I used a between-subjects design in this study, while also recognizing that this choice came with trade-offs.

Despite these limitations, there are several strengths of Study 3. In particular, the tightly controlled, between-subjects experimental design of Study 3 enabled me to isolate the processes underlying predicted and/or actual discrete versus dimensional states and examine whether the resulting forecasting errors differed. Additionally, by norming stimuli for individual discrete emotions, I was able to induce the expectation and experience of specific emotions and compare forecasting errors for these emotions to dimensional states for the same stimuli. Lastly, by a utilizing a new discrete approach (i.e., by asking people to rate one specific discrete emotion) and a dimensional approach not used in Study 1 (i.e., by asking people to directly rate how negative/positive they would feel/felt), I was able to examine whether affective forecasting errors differed when they were measured using discrete versus dimensional measures from a new angle.

General Discussion

In this dissertation, I described four studies: Study 1 was a within-subjects study; Study 3 was a between-subjects study; Studies 2a and 2b validated the stimuli used in Study 3. The overall aim was to examine whether affective forecasting errors differed based on the type of measurement approach used to study them.

Advances in Understanding

Existing research on affective forecasting has focused on the idea that people make affective forecasting errors about the intensity of their future emotions. However, some evidence suggests that this is not always the case, and that people can also make accurate affective forecasts. In this research I examined one novel factor that could explain some of the variation in these findings – the divergent measurement approaches that affective forecasting researchers use. Affective forecasting researchers either use discrete or dimensional approaches to measure and analyze affective forecasting errors. Virtually no studies to date have used both approaches in a single study to compare affective forecasting errors for discrete and dimensional states. Furthermore, a majority of the existing affective forecasting research has examined affective forecasts about events that likely induce high arousal positive and negative emotions such as presidential elections (Barber et al., 2023), romantic breakups (Eastwick et al., 2008), and tenure decisions (Gilbert et al., 1998). There is a lack of studies examining affective forecasting errors for lower arousal positive and negative emotions. In this work, I conducted two studies that enabled me to compare affective forecasting errors between measurement types using pictures that induced emotions for both higher and lower arousal positive and negative emotions.

Contrary to our hypothesis, results from both studies suggested that affective forecasting errors do not differ based on measurement approaches, at least in the context of the pictures that were used in Studies 1 and 3. While these null results could be attributed in part to the methodological limitations of these studies, it is also likely that in reality, the hypothesis is incorrect, and that affective forecasting errors do not actually differ based on measurement approaches. This finding has both practical and theoretical implications for the literature on affective forecasting.

Practically, if affective forecasting errors do not differ as a function of discrete and dimensional approaches, this suggests that forecasting researchers might be able to use these approaches interchangeably, depending on the context of their studies. Existing affective forecasting studies typically employ either the discrete (e.g. Sekhsaria and Pronin, 2021) or the dimensional approach (e.g., Holloway and Weiner, 2021) depending on their research questions and the context of their study. These findings suggest that because affective forecasting errors as measured by discrete and dimensional approaches ultimately lead to the same findings, researchers can choose the approach that is best suited for their research question without needing to worry about their findings being biased by their choice of measurement approach. As a result, they could have more confidence in the effects of their experimental manipulations on affective forecasting errors.

That is, if a researcher is specifically interested in studying affective forecasting errors for the discrete emotion of disgust, they can measure disgust directly, as opposed to needing to measure negative affect. Since our findings suggest that affective forecasting errors do not differ based on measurement approach, if this researcher does uncover evidence of forecasting errors, they can be confident that the magnitude of error is *not* biased by their measurement approaches (i.e., that they would have likely obtained the same result using the dimensional approach).

Theoretically, the implications are not as straightforward. Despite the evidence from these studies suggesting affective forecasting errors do not differ based on measurement approaches, this does not necessarily imply that the processes underlying discrete and dimensional affective forecasts or ratings of actual affect are also similar. The processes underlying forecasts and/or ratings for actual affect *could still be distinct* even though they manifest as similar levels of affective forecasting errors. Alternatively, it is also possible, given that these studies found that affective forecasting errors did not differ based on measurement approaches, that the processes underlying these forecasts and ratings for discrete and dimensional states are actually the same. Overall, it is difficult to commit to just one of these implications on the broader affective forecasting literature until future researchers: 1) test whether affective forecasting errors differ based on measurement approaches in a context different from the one I used (i.e., use different stimuli), and 2) explicitly measure the processes underlying discrete and dimensional affective forecasting errors.

Broadly, the present studies fill gaps in the affective forecasting literature in the following ways. First, these studies filled the gap in the existing research by using both discrete and dimensional measurement approaches to compare affective forecasting errors. Second, Study 1 used one way of using the discrete approach (i.e., by measuring several discrete emotion items) and one way of using the dimensional approach (i.e., by asking participants to make ratings about valence and arousal); Study 3 adopted a different conceptualization of the both the discrete (i.e., by asking participants to predict and rate individual primary discrete emotions) and dimensional (i.e., by asking participants to directly predict how positive/negative they would feel) approaches. Thus, the studies collectively provided a comprehensive understanding of whether affective forecasting errors differ when they are measured using discrete versus dimensional approaches. Third, given the lack of studies examining affective forecasting errors for low and high arousal positive and negative emotions, both studies filled this gap by examining affective forecasting errors for a variety of discrete and dimensional emotions/affective states that span every quadrant of the affective circumplex. Lastly, while Study 1 provided an understanding of whether affective forecasting errors differ using discrete and dimensional approaches in a within-subjects experimental context, Study 3 tested this idea using a between-subjects experimental context.

Limitations and Directions for Future Research

There are several promising directions for future research suggested by this work. While results suggested that affective forecasting errors likely do not differ as a function of measurement approach, there are limitations spanning the two main studies that could be worth addressing in future research.

First, both the main studies in this research, Studies 1 and 3, tested affective forecasting errors in the context of emotionally evocative, validated pictures and picture descriptions. Past research on affective forecasting has focused on examining high arousal emotions such as fear, anxiety, and excitement (Chanel et al., 2022). I selected picture stimuli as opposed to other forms of stimuli (at least in Study 3) so as to move beyond the well-studied, high arousal emotions in the existing literature and elicit the experience and expectation of emotions that span *every quadrant* of the entire affective circumplex. While pictures have been shown to reliably induce emotions in laboratory settings (Lang et al., 2005), they have one major drawback: they might not always provide the most ecologically valid experience of emotions. That is, experiencing an emotion in the real world is often a more dynamic process than viewing a single emotionally evocative picture for 12 seconds, for example. It often involves multiple components such as changes in one's physiology, subjective experiences, and overt behaviors (Matsumoto & Hwang, 2012; Romeo et al., 2022). Research suggests that using dynamic stimuli, such as short video clips, might be one of the more effective techniques for studying emotional experiences and processing in a laboratory setting (Rottenberg et al., 2007, Romeo et al., 2022). Furthermore, dynamic stimuli are more ecologically valid because they provide a participant with both visual and auditory information, thus making participants' lab-based experience parallel their experience in reality (Romeo et al., 2022).

I recommend that future researchers replicate these studies with video clips instead of pictures. It is possible that making predictions about emotions/affect based on descriptions of videos induces more vivid simulations, stronger premotions, and provides a different context compared to those that are induced from descriptions of pictures. Similarly, it's possible that videos provide more realistic situations that demand participants' attention, induce appraisal of relevance, and thus stronger emotions than pictures do. It is possible that there is an effect of measurement type on affective forecasting errors, but because the experience of emotions in this study was not intense enough and might not have paralleled people's experiences in the real world, I was unable to detect this effect. Future researchers could benefit from using videos instead as this might induce the expectation and experience stronger and more realistic emotions, potentially more variation in affective forecasting errors and allow them to uncover the effects of measurement type, should they exist.

To further enhance ecological validity, researchers could also replicate this research using a daily life method. In particular, daily life methods enable researchers to examine emotions that are more vivid, personally relevant, and impactful as compared to those that are elicited in a laboratory (Kuppens et al., 2022; Moeck et al., 2022). Furthermore, this method allows for a nuanced examination of the variation and complexity of the emotions people experience in their daily lives (Kuppens et al., 2022). Future researchers could, for example, ask participants to nominate the most positive and negative event for the next day, and predict which emotions and affective states they will predominantly feel, and how much of these emotions they will feel. It is possible that making predictions about emotions/affect for highly self-relevant events induces more vivid simulations, stronger premotions, and provides a different context compared to those that are induced from picture descriptions in a laboratory. Then, at the end of the next day, participants could be asked to rate how much they actually experienced those emotions and affective states. Similarly, it’s possible that experiencing these events firsthand provides a more realistic situation that arguably demands more of participants’ attention, induces appraisal of relevance, and thus stronger emotions than simply viewing pictures in the laboratory. Overall, using daily life methods might increase the self-relevance for participants, induce the expectation and experience of stronger and more realistic emotions, potentially more variation in affective forecasting errors, and reveal any effects of measurement type on forecasting errors, should they exist.

Second, even though I did not find an effect of measurement type on affective forecasting errors, I still believe that it is worth examining the processes underlying affective forecasting errors for discrete and dimensional measures. As detailed in the previous section, it is possible that even though discrete and dimensional measures led to similar levels of affective forecasting errors, the processes occurring “under the hood” might differ. Researchers could test this idea by using a qualitative approach in their affective forecasting studies. Specifically, they could ask participants to make affective forecasts using discrete and/or dimensional measures (depending on their design) and subsequently administer an open-ended thought-listing measure (sometimes known as a retrospective report) to them. This measure could ask participants to describe every thought or idea that they had while they made their affective forecasts and rated their actual affect (Buehler & McFarland, 2001; Patel & Urry, in preparation). Researchers could then code these responses for themes and examine whether people's thoughts, simulations, and premotions at the time of forecasting and actual affect ratings differ when measured using discrete and dimensional approaches. In addition to providing a nuanced understanding of affective forecasting errors, research that tests this idea would also contribute to the sparse literature on the processes underlying affective forecasting errors in general (Patel & Urry, in preparation; Gilbert & Wilson, 2007; 2009).

Third, this research examined *one* potential reason for mixed findings in the affective forecasting literature. Since I did not find support for the idea that measurement approaches account for some of the variation in findings in this area of research, I recommend that future researchers re-consider sources of variation and examine other potential reasons for mixed findings. One such source of variation that might be worth examining is the nature of the event that participants are asked to make forecasts about.

In particular, a majority of the studies in the existing affective forecasting literature ask people to create their affective forecasts about the *outcome* of future events such as not being awarded tenure (Gilbert et al., 1998; Wilson et al., 2000; Hoerger et al., 2016; Hezel et al., 2019; Verner-Fillion et al., 2012). However, some, although fewer studies have also measured affective forecasts about how people will feel *during* upcoming events such as an upcoming race (Aitken et al., 2023; Zelenski et al., 2013; Nielsen et al., 2008; Kong, 2015; Mallet et al., 2008). The literature on mental simulations suggests that simulating the outcome and the process of future events are distinct phenomena that have different affective consequences (Taylor et al., 1998; Pham & Taylor, 1999). Since one of the ways in which people make their affective forecasts is by mentally simulating events (Gilbert & Wilson, 2007; 2009), it is likely that making affective forecasts about how one will feel about the *outcome* of a situation versus the *process of being in* a situation constitute empirically distinct phenomena that subsequently result in different levels of affective forecasting errors.

Given this variation in the nature of the events that researchers examine, it is possible that affective forecasting errors differ *systematically* based on whether they are measured in reference to the process or outcome of an event. To date, one study (Aitken et al., 2023) examined forecasting errors about the process of running a race *and* being finished with the race. These researchers found that participants made affective forecasting errors for both running and completing the race. However, these results are from just one study and one context (i.e., a race). Future researchers could benefit from examining whether affective forecasting errors for the process and outcome of an event differ in another context and perhaps using a meta-analysis of existing studies to make sense of the mixed findings in this area of research.

Fourth, while the studies presented here provided experimental evidence that suggested that affective forecasting errors may not differ based on measurement approaches, I believe that it may also behoove future researchers to conduct a systematic meta-analysis of the existing research to verify these findings. Given that a recent review of the literature revealed that there is reasonable variation in measurement approaches in forecasting studies (Patel & Urry, 2024) and that the last systematic meta-analysis on affective forecasting errors was conducted in 2012 (Mathieu & Gosling, 2012), I believe that such an endeavor could help advance the field. This will allow researchers to have a comprehensive understanding of whether measurement approaches can account for some of the variation in mixed findings in this area of research using a variety of data sources and experimental and non-experimental designs.

Specifically, a meta-analysis could determine whether more studies use one type of measurement approach to affective forecasting errors over another, and whether the direction or magnitude of affective forecasting errors differs for dimensional versus discrete emotions/affective states. If such a meta-analysis reveals that errors for forecasts of affect are larger or smaller than those for discrete emotions, this could suggest that forecasting errors depend, in part, on the discrete versus dimensional state being forecast and, thus, are not exclusively marking trait-like differences in overall forecasting abilities. Additionally, it would encourage researchers to be more mindful of the measurements they collect and the inferences they draw.

Fifth, in this research I examined the *magnitude* and *direction* (Studies 1 and 3) of affective forecasting errors based on discrete and dimensional approaches. However, another feature of affective forecasting errors is their consequences. Past research indicates that there is a link between the direction of affective forecasting errors and outcomes such as task performance. For example, Kaplan and colleagues (2020) found that when people underpredicted their positive affect (predicted less positive affect than they experienced) and when they overpredicted their negative affect (predicted greater negative affect than they experienced), they reported better task performance. While these studies examined whether the magnitude and direction of affective forecasting errors differed based on measurement approaches, given past research such as Kaplan's (2020), it may be worth extending the existing research and examining whether affective forecasting errors measured using discrete and dimensional approaches also have unique downstream consequences on performance outcomes.

Lastly, in this research I explored several moderators of the association between measurement approaches and affective forecasting errors such as baseline affect, age, stimulus valence, and stimulus arousal. While this was a sufficient list of moderators, given the study designs, there might still be other hidden moderators of the relationship between measurement approaches and affective forecasting errors that I did not capture in these studies. Including these moderators could account for part of the variation that is being considered as noise in statistical analyses, and thereby increase the sensitivity of statistical models to detect any effects of type of affect rating and measurement type.

One such moderator that could be interesting for future researchers to examine is that of metacognitive ability. Broadly speaking, metacognitive ability is the ability to reflect upon one's thoughts and behavior and to have an awareness of one's cognitive processes (Metcalf, 1994). One could argue that when participants are asked to make affective forecasts about individual discrete emotions, this inadvertently introduces a procedural confound, as it provides and/or *requires* them to have higher metacognitive abilities in order to make ratings using fine-grained discrete emotion categories. This is in contrast to participants who are asked to make forecasts about how broadly negative/positive they will feel, as these measures do not necessarily require high levels of metacognitive abilities. As a result, metacognitive abilities could moderate the association between measurement approaches and affective forecasting errors.

For example, participants who have higher metacognitive abilities might make similar affective forecasting errors for discrete and dimensional states. This could be because they are able to identify, monitor, predict, and report both discrete and dimensional states equally accurately because of their inherently high metacognitive abilities. On the other hand, participants who have lower metacognitive abilities might exhibit larger affective forecasting errors for discrete rather than dimensional states. This could be because they might find it harder to identify, monitor, predict, and report individual discrete emotions, but have less difficulty doing so for broader, general affective states. I did not measure metacognitive ability in Study 3 as the random assignment would have accounted for participants' unique levels of metacognitive abilities. However, I recommend that future researchers replicate our studies and examine the role of metacognitive abilities as a moderator of the relationship between measurement type and affective forecasting errors.

Relatedly, a growing body of literature suggests that high metacognitive abilities are linked to a host of behavioral outcomes. Specifically, people with higher metacognitive abilities exhibit improved reading comprehension, problem-solving, decision making (e.g., Thompson & Johnson, 2014), and overall academic success (e.g., Young & Fry, 2008). They are also better able to regulate their negative emotions and exhibit higher levels of well-being than those with lower metacognitive abilities (e.g., Temircan, 2023). Lastly, these individuals also have better interpersonal skills and social relationships (e.g., Jain & Imran, 2024). Thus, if affective forecasting errors do differ based on measurement approaches and individuals' metacognitive abilities, this might ultimately lead to unique effects on participants' learning, academic success, decision-making, mental health, and social relationships. Thus, future research is needed to fully understand the relationship between affective forecasting errors as measured by discrete and dimensional approaches, constructs such as metacognitive abilities, and downstream effects of these associations.

Concluding Comment

Overall, the research presented in this dissertation aimed to better understand one possible reason for mixed findings in the affective forecasting area. These studies: 1) utilized both discrete and dimensional approaches to measuring affective forecasting errors, 2) examined whether affective forecasting errors differ between measurement approaches in two tightly controlled experimental settings, and 3) investigated whether affective forecasting errors for discrete emotions and dimensional affective states differed for a variety of low and high arousal positive and negative emotions/affective states. Across studies, results suggested that affective forecasting errors do not differ based on measurement approaches, at least in the context of pictures that were used. Nevertheless, these results are from one set of studies; future research should build on and extend this work to uncover whether these results replicate or whether there is in fact a difference between affective forecasting errors based on measurement approaches. Overall, these studies extend the research on affective forecasting, examine a novel source of variation in findings, and provide a nuanced understanding of affective forecasting errors.

References

Aitken, J. A., Kaplan, S. A., Pagan, O., Wong, C. M., Sikorski, E., & Helton, W. (2021). Affective Forecasts for the Experience Itself: An Investigation of the Impact Bias during an Affective Experience. *Current Psychology*. https://doi.org/10.1007/s12144-021-02337-8

Barber, S. J., Kausar, H., & Udry, J. (2023). Age differences in affective forecasting accuracy. *Psychology and Aging*. https://doi.org/10.1037/pag0000722

Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition and Emotion, 12*(4), 579 – 599.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01

Chanel, O., Lyk-Jensen, S. V., and Vergnaud, J. (2022). Does affective forecasting error induce changes in preferences? Lessons from Danish soldiers anticipating combat in Afghanistan. *Defence and Peace Economics, 34,* 660–683. doi: 10.1080/ 10242694.2022.2037829

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* New York, NY: Routledge Academic.

Coundouris, S. P., Henry, J. D., Suddendorf, T., & Lehn, A. C. (2023). Affective forecasting in Parkinson's disease. *Journal of the International Neuropsychological Society: JINS, 29*(4), 406–409. https://doi.org/10.1017/S1355617722000388

Dorinson, C. A., Minson, J. A., & Rogers, T. (2019). Selective exposure partly relies on faulty affective forecasts. Cognition, 188, 98-107. https://doi.org/10.1016/j.cognition.2019.02.010

Dunn, E. W., Brackett, M. A., Ashton-James, C., Schneiderman, E., & Salovey, P. (2007). On Emotionally Intelligent Time Travel: Individual Differences in Affective Forecasting Ability. *Personality and Social Psychology Bulletin*, *33*(1), 85–93. https://doi.org/10.1177/0146167206294201

Eastwick, P. W., Finkel, E. J., Krishnamurti, T., & Loewenstein, G. (2008). Mispredicting distress following romantic breakup: Revealing the time course of the affective forecasting error. *Journal of Experimental Social Psychology, 44*(3), 800–807. https://doi.org/10.1016/j.jesp.2007.07.001

Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2)*,* 124–129. https://doi.org/10.1037/h0030377

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Floerke, V. A., & Urry, H. L. Harris, L. (2020). *Calibrating Optimal Forecasting for Emotional Events* [Unpublished doctoral dissertation]. Tufts University.

Frank, C. C., Iordan, A. D., Ballouz, T. L., Mikels, J. A., & Reuter-Lorenz, P. A. (2021). Affective forecasting: A selective relationship with working memory for emotion. *Journal of Experimental Psychology: General, 150*(1), 67–82. https://doi.org/10.1037/xge0000780

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*(3), 617–638. https://doi.org/10.1037/0022-3514.75.3.617

Gilbert, D. T., & Wilson, T. D. (2007). *Prospection: Experiencing the future. Science, 317*(5843), 1351–1354. https://doi.org/10.1126/science.1144161

Gilbert, D. T., & Wilson, T. D. (2009). Why the brain talks to itself: Sources of error in emotional prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1521), 1335–1341. https://doi.org/10.1098/rstb.2008.0305

Gross, J. J., & Thompson, R. A. (2007). Emotion Regulation: Conceptual Foundations. In J. J. Gross (Ed.), *Handbook of Emotion Regulation* (pp. 3–24). The Guilford Press.

Gross, J. J., Sheppes, G., & Urry, H. L. (2011). Cognition and Emotion Lecture at the 2010 SPSP Emotion Preconference: Emotion generation and emotion regulation: A distinction we should make (carefully). *Cognition and Emotion, 25*(5), 765–781. https://doi.org/10.1080/02699931.2011.555753

Hezel, D. M., Stewart, S. E., Riemann, B. C., & McNally, R. J. (2019). Affective forecasting accuracy in obsessive compulsive disorder. *Behavioural and Cognitive Psychotherapy, 47(*5), 573–584. https://doi.org/10.1017/S1352465819000134

Hoerger, M., Scherer, L. D., & Fagerlin, A. (2016). Affective Forecasting and Medication Decision Making in Breast Cancer Prevention. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *35*(6), 594–603. https://doi.org/10.1037/hea0000324

Hoerger, M., Chapman, B., & Duberstein, P. (2016). Realistic affective forecasting: The role of personality. *Cognition & Emotion, 30*(7), 1304–1316. https://doi.org/10.1080/02699931.2015.1061481

Holloway, C. P. & Wiener, R. L. (2021). Affective forecasting and ex-offender hiring decisions. *Motivation and Emotion, 17*(45), 489-505. doi: 10.1007/s11031-021-09885-3

Jain, A. & Imran, M. (2024). Influence of metacognition on social competence and emotional maturity in young adults. *International Journal of Indian Psychology, 12*(2), 1444-1453. doi:18.01.126.20241202, DOI:10.25215/1202.126

Kaplan, S., Winslow, C., Craig, L., Lei, X., Wong, C., Bradley-Geist, J., Biskup, M., & Ruark, G. (2020). "Worse than I anticipated" or "This isn't so bad"?: The impact of affective forecasting accuracy on self-reported task performance. *PLOS ONE*, *15*(7), e0235973. https://doi.org/10.1371/journal.pone.0235973

Karl, M., Kock F., Ritchie, B. W., & Gauss, J. (2021). Affective forecasting and travel decision-making: An investigation in times of a pandemic. *Annals of Tourism Research, 87,* 103139. https://doi.org/10.1016/j.annals.2021.103139

Kong, D. T. (2015). The role of mindfulness and neuroticism in predicting acculturative anxiety forecasting error. *Mindfulness, 6*(6), 1387–1400. https://doi.org/10.1007/s12671-015-0409-4

Kuppens, P., Dejonckheere, E., Kalokerinos, E. K., & Koval, P. (2022). Some recommendations on the use of daily life methods in Affective Science. *Affective Science, 3,* 505–515. https://doi.org/10.1007/s42761-022-00101-0

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Instruction manual and affective ratings, Technical Report A-8.* Gainesville: The Center for Research in Psychophysiology, University of Florida.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*(12), 1248–1263. https://doi.org/10.1016/S0006-3223(98)00275-3

Lench, H. C., Levine, L. J., Perez, K., Carpenter, Z. K., Carlson, S. J., Bench, S. W., & Wan, Y. (2019). When and why people misestimate future feelings: Identifying strengths and weaknesses in affective forecasting. *Journal of Personality and Social Psychology*, *116*(5), 724–742. https://doi.org/10.1037/pspa0000143

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*(2), 278–285. https://doi.org/10.1037/a0022792

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*(4), 584–605. https://doi.org/10.1037/a0029544

Mallett, R. K., Wilson, T. D., & Gilbert, D. T. (2008). Expect the unexpected: Failure to anticipate similarities leads to an intergroup forecasting error. *Journal of Personality and Social Psychology, 94*(2), 265–277. https://doi.org/10.1037/0022-3514.94.2.94.2.265

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2013). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods, 46,* 596 - 610. doi:10.3758/s13428-013-0379-1

Mathieu, M., & Gosling, S. (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

Matsumoto, D., and Hwang, H. S. (2012). Culture and emotion: the integration of biological and cultural contributions. *Journal of Cross-Cultural Psychology, 43*, 91–118. doi: 10.1177/0022022111420147

Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about knowing.* The MIT Press. https://doi.org/10.7551/mitpress/4561.001.0001

Moeck, E., Grewal, K., Greenaway, K. H., Koval, P., & Kalokerinos, E. K. (2022). Affective forecasting in everyday life: Accuracy and associations with emotional benefits. https://doi.org/10.31234/osf.io/sr9vj

Nielsen, L., Knutson, B., & Carstensen, L. L. (2008). Affect dynamics, affective forecasting, and aging. Emotion, 8(3), 318–330. https://doi.org/10.1037/1528-3542.8.3.318

Öhman, A., & Mineka, S. (2003). The malicious serpent: Snakes as a prototypical stimulus for an evolved module of fear. *Current Directions in Psychological Science, 12*(1), 5-9. https://doi.org/10.1111/1467-8721.01211

OpenAI. (2023). ChatGPT (Oct 2023 version) [Large language model]. OpenAI. https://openai.com

O'Toole, M. S., Renna, M. E., Elkjær, E., Mikkelsen, M. B., and Mennin, D. S. (2020). A systematic review and meta-analysis of the association between complexity of emotion experience and behavioral adaptation. *Emotion Review, 12,* 23–38. doi: 10.1177/1754073919876019

Patel, P., & Urry, H. L. (2024). Discrete and dimensional approaches to affective forecasting. *Frontiers in Psychology, 15,* 1412398. https://doi.org/10.3389/fpsyg.2024.1412398

Patel, P., & Urry, H. L. *Everything is going to be okay: Emotion Regulation and Immune Neglect in Affective Forecasting.* Manuscript under review.

Patel, P., & Urry, H.L. *Individual differences in affective forecasting under stress.* Manuscript in preparation.

Petagna, K. D., & Wormwood, J. B. (2023). Who can predict their future feelings? Individual differences in affective forecasting accuracy. *Social Psychological and Personality Science*, *0*(0). https://doi.org/10.1177/19485506231208749

Pham, L. B., & Taylor, S. E. (1999). From thought to action: Effects of process-versus outcome-based mental simulations on performance. *Personality and Social Psychology Bulletin, 25*(2), 250-260. https://doi.org/10.1177/0146167299025002010

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734. https://doi.org/10.1017/S0954579405050340

Romeo, Z., Fusina, F., Semenzato, L., Bonato, M., Angrilli, A., & Spironelli, C. (2022). Comparison of slides and video clips as different methods for inducing emotions: An electroencephalographic alpha modulation study. *Frontiers in Human Neuroscience, 16,* 901422. https://doi.org/10.3389/fnhum.2022.901422

Rottenberg, J., Ray, R. D., and Gross, J. J. (2007). "Emotion elicitation using films," in *The Handbook of Emotion Elicitation and Assessment*, eds J. A. Coan and J. J. B. Allen (New York, NY: Oxford. University).

Ruby, M. B., Dunn, E. W., Perrino, A., Gillis, R., & Viel, S. (2011). The invisible benefits of exercise. *Health Psychology, 30*(1), 67–74. https://doi.org/10.1037/a0021859

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*(5), 805–819. https://doi.org/10.1037/0022-3514.76.5.805

Sekhsaria, S., and Pronin, E. (2021). Underappreciated benefits of Reading own and others' memories. *Social Cognition, 39,* 504–525. doi: 10.1521/soco.2021.39.4.504

Scheibe, S., Mata, R., & Carstensen, L. L. (2011). Age differences in affective forecasting and experienced emotion surrounding the 2008 US presidential election. *Cognition and Emotion, 25*(6), 1029–1044. https://doi.org/10.1080/02699931.2010.545543

Tabachnick, B. G., & Fidell, L. S. (2018). *Using multivariate statistics (7th ed.).* Pearson.

Takano, K., & Ehring, T. (2023). Affective forecasting as an adaptive learning process. *Emotion.* Advance online publication. https://doi.org/10.1037/emo0001303

Taylor, S. E., Pham, L. B., Rivkin, I. D., & Armor, D. A. (1998). Harnessing the imagination: Mental simulation, self-regulation, and coping. *American Psychologist, 53*(4), 429–439. https://doi.org/10.1037/0003-066X.53.4.429

Temircan, Z. (2023). Exploring the relationship between metacognition, emotional regulation and perceived stress among college students. *Current Approaches in Psychiatry, 15,* 110-118. doi: 10.18863/pgy.1246718

Thompson, R. J., Spectre, A., Insel, P., Mennin, D., Gotlib, I. H., & Gruber, J. (2017). Positive and negative affective forecasting in remitted individuals with Bipolar I Disorder, and Major Depressive Disorder, and healthy controls. *Cognitive Therapy and Research, 41*(5), 673–685. https://doi.org/10.1007/s10608-017-9840-2

Thompson, V. A., & Johnson, D. M. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning, 20*(2), 215-244. https://doi.org/10.1080/13546783.2013.869763

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 114–137). The Guilford Press.

Urry, H. L., & Gross, J. J. (2010). Emotion regulation in older age. *Current Directions in Psychological Science, 19*(6), 352–357. https://doi.org/10.1177/0963721410388395

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219–235. https://doi.org/10.1037/0033-2909.98.2.219

Verner-Filion, J., Lafrenière, M. K., & Vallerand, R. J. (2012). On the accuracy of affective Forecasting: The moderating role of passion. *Personality and Individual Differences,* *52*(7), 849–854. https://doi.org/10.1016/j.paid.2012.01.014

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*, 820–838. https://doi.org/10.1037/0022-3514.76.5.820

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

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*(5), 821–836. https://doi.org/10.1037//0022-3514.78.5.821

Yik, M., Russell, J. A., & Steiger, J. H. (2011). A 12-Point Circumplex Structure of Core Affect. *Emotion*, *11*(4), 705–731. https://doi.org/10.1037/a0023980

Young, A., & Fry, J. D. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning, 8*(2), 1-10.

Zelenski, J. M., Whelan, D. C., Nealis, L. J., Besner, C. M., Santoro, M. S., & Wynn, J. E. (2013). Personality and affective forecasting: Trait introverts underpredict the hedonic benefits of acting extraverted. *Journal of Personality and Social Psychology*, *104*(6), 1092–1108. https://doi.org/10.1037/a0032281

Supplementary Materials

Study 1

Table S1

Results from linear mixed effects regression on absolute affective forecasting error scores, including fixed effects for measurement type (deviation-coded), and random intercepts for participants, emotions, and stimulus number for Study 1

	Absolute affective forecasting error scores		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	0.93	0.79 – 1.08	<0.001***
Dimensional-Discrete	-0.03	-0.28 – 0.22	0.801
Random Effects			
σ^2		1.57	
$\tau_{00\ ID}$		0.08	
$\tau_{00\ stimNum}$		0.02	
$\tau_{00\ emotion}$		0.03	
ICC		0.07	
N_{ID}		125	
$N_{emotion}$		17	
$N_{stimNum}$		20	
Observations		42500	
Marginal R^2 / Conditional R^2		0.000 / 0.073	

Note. *p < .05; **p <.01; ***p < .001

Study 1

Table S2

Results from linear mixed effects regression on responses to affect ratings, including fixed effects for type of affect rating (predicted, actual), measurement type (dimensional, discrete), stimulus valence (pleasant, unpleasant), and stimulus arousal (low, high) (all deviation-coded), and random intercepts for participants, emotions, and stimulus number for Study 1

	Affect rating		
Predictors	*Estimates*	*95%CI*	*p*
Intercept	3.41	3.05 – 3.77	<0.001***
Predicted-Actual	0.02	-0.02 – 0.06	0.368
Dimensional-Discrete	1.70	1.02 – 2.37	<0.001***
Stimulus valence: Pleasant-Unpleasant	0.83	0.65 – 1.00	<0.001***
Stimulus arousal: Low-High	-0.14	-0.32 – 0.03	0.108
Predicted-Actual × Dimensional-Discrete	0.05	-0.03 – 0.13	0.199
Predicted-Actual × Pleasant-Unpleasant	-0.07	-0.14 – 0.01	0.085
Dimensional-Discrete × Stimulus valence: Pleasant-Unpleasant	1.48	1.41 – 1.56	<0.001***
Predicted-Actual × Stimulus arousal: Low-High	0.09	0.02 – 0.17	0.019*
Dimensional-Discrete × Stimulus arousal: Low-High	0.49	0.42 – 0.57	<0.001***
Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	0.40	0.05 – 0.74	0.025*

Predicted-Actual × Dimensional-Discrete × Stimulus valence: Pleasant-Unpleasant	-0.10	-0.25 – 0.06	0.216
Predicted-Actual × Dimensional-Discrete × Stimulus arousal: Low-High	0.13	-0.02 – 0.28	0.094
Predicted-Actual × Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	0.00	-0.15 – 0.15	0.997
Dimensional-Discrete × Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	-0.53	-0.68 – -0.37	<0.001***
Predicted-Actual × Dimensional-Discrete × Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	0.04	-0.27 – 0.34	0.820
	Random Effects		
σ^2		3.36	
$\tau_{00\ ID}$		0.27	
$\tau_{00\ stimNum}$		0.04	
$\tau_{00\ emotion}$		0.21	
ICC		0.13	
N_{ID}		126	
$N_{emotion}$		17	
$N_{stimNum}$		20	
Observations		85340	
Marginal R^2 / Conditional R^2		0.100 / 0.220	

Note. **p* < .05; ***p* <.01; ****p* < .001

Table S3

Results from linear mixed effects regression on affective forecasting error scores, including fixed effects for emotion valence and type of affect rating (deviation-coded), and random intercepts for participants, and stimulus number for Study 1

	Affect rating		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	3.41	3.18 – 3.64	<0.001***
Predicted-Actual	0.02	-0.02 – 0.06	0.350
Arousal-Valence	1.15	1.11 – 1.18	<0.001***
Discrete Negative - Valence	-1.05	-1.09 – -1.02	<0.001***
Discrete Positive - Valence	-0.63	-0.67 – -0.59	<0.001***
Predicted-Actual × Arousal-Valence	0.00	-0.07 – 0.07	0.971
Predicted-Actual × Discrete Negative - Valence	-0.10	-0.17 – -0.03	0.008**
Predicted-Actual × Discrete Positive - Valence	0.05	-0.02 – 0.13	0.152
Random Effects			
σ^2	2.32		
$\tau_{00\ ID}$	0.11		
$\tau_{00\ stimNum}$	0.25		
ICC	0.14		
N_{ID}	126		
$N_{stimNum}$	20		
Observations	20080		
Marginal R^2 / Conditional R^2	0.225 / 0.330		

Note. *p < .05; **p <.01; ***p < .001

Study 1: Descriptions for the Affective Forecasting Task

1. A baby seal with a fluffy white coat in the snow
2. A young antelope drinking water from a stream
3. A young girl with a flower in her hair smiling
4. A toddler hugging and kissing a puppy's cheek
5. A middle-aged couple riding their bikes together smiling.
6. A woman sitting by herself at a table for two people, staring into the distance and frowning
7. A close-up photo of a person getting fingerprinted
8. An older man holding a pot full of food in a kitchen
9. A person locked in a jail cell holding a cigarette through the bars
10. A mop and a mop bucket filled with brown dirty water
11. A group of people doing synchronized skydiving, holding hands in a circle while free falling above the water
12. A skier at the top of a steep long-distance ski jump
13. An Olympic gymnast raising her hands above her head in triumph
14. People on a moving roller coaster raising their hands above their head while screaming and smiling
15. A large pile of money with various bills, including fifty-, twenty-, five- and one-dollar bills.
16. A snake hanging off a branch, posed to strike with its jaws open
17. A man driving a car while holding an open bottle of beer
18. A massive dark tornado with a city in front of its path.

19. A man holding a gun up to the side of his head with his finger on the trigger

20. A ship sinking into the ocean with more than half of it submerged under the water

Study 2a

Table S4

Sample demographics for Studies 2a and 2b

	Study 2a	Study 2b
Mean age in years (*SD*)	38(*12*)	40.10(*12.72*)
Gender		
Male	34%	33.33%
Female	58%	66.67%
Non-binary	6%	-
Other	2%	-
Ethnicity/Race		
White	54%	64.58%
African American	26%	14.58%
Asian/Pacific Islander	10%	6.25%
Native American	-	-
Latino/a/x	4%	8.33%
Multi-racial	6%	6.25%
Other	-	-
Decline	-	-
Education Level		
Some high school	2%	-
High school (diploma, GED, or alternative credential)	12%	12.50%
Some college credit, but less than 1 year of college credit	4%	6.25%
1 or more years of college credit, no degree	20%	6.25%
Associate’s degree (for example: AA, AS)	6%	14.58%
Bachelor’s degree (for example: BA, BS)	34%	41.67%
Master’s degree (for example: MA, MS, MEng, MEd, MSW, MBA)	20%	12.50%
Professional degree beyond a bachelor’s degree	2%	2.08%

(for example: MD, DDS, DVM, LLB, JD)		
Doctorate degree (for example: PhD, EdD)	-	4.17%
Political Ideology		
Very liberal	18%	10.42%
Liberal	26%	25%
Somewhat liberal	6%	10.42%
Moderate	26%	18.75%
Somewhat Conservative	6%	6.25%
Conservative	18%	18.75%
Very Conservative	-	10.42%
Color Blindness		
Yes	-	-
No	98%	100%
Not sure	2%	-

Table S5

Information about the picture stimuli used in Studies 2a and 2b. Bolded rows were picture stimuli selected for Study 2b

Discrete Emotion	Picture Number	Content of Picture	Source	Picture Number from Source	Link (if any)
Anger	Picture 1	School shooter	NAPS	Faces_270_h	N/A
Anger	**Picture 2**	**People beating up a man**	**NAPS**	**People_127_h**	N/A
Anger	Picture 3	Angry man	Google	N/A	https://www.flickr.com/photos/theimagegroup/322557849/in/photostream/
Sadness	**Picture 1**	**Old thin man in bed**	**NAPS**	**Faces_172_h**	N/A
Sadness	Picture 2	Homeless person sleeping	NAPS	People_143_h	N/A
Sadness	Picture 3	Two burn victims	NAPS	People_140_h	N/A
Fear	Picture 1	Man with gun	IAPS	6250	N/A
Fear	**Picture 2**	**Wolf snarling**	**NAPS**	**Animals_004_v**	N/A
Fear	Picture 3	Woman being attacked with knife	IAPS	6550	N/A
Happiness	Picture 1	Dog with stick at the beach	NAPS	Animals_183_h	N/A
Happiness	**Picture 2**	**Child laughing in the park**	**NAPS**	**Faces_122_h**	N/A
Happiness	Picture 3	Children laughing	IAPS	2347	N/A
Disgust	Picture 1	Mangled hand	IAPS	3150	N/A
Disgust	Picture 2	Man vomiting	IAPS	9325	N/A

Discrete Emotion	Picture Number	Content of Picture	Source	Picture Number from Source	Link (if any)
Disgust	**Picture 3**	**Dirty toilet**	**NAPS**	**Objects_125_h**	N/A
Fatigue	**Picture 1**	**Athlete hunched over**	**IAPS**	**8121**	N/A
Fatigue	Picture 2	Old man walking in the snow	NAPS	People_138_h	N/A
Fatigue	Picture 3	Man and dog asleep	NAPS	People_149_h	N/A
Alert	**Picture 1**	**Bustling city**	**Google**	**N/A**	https://www.flickr.com/photos/kevinpoh/8556961493
Alert	Picture 2	Vibrant colors	Google	N/A	https://www.goodfon.com/abstraction/wallpaper-background-rainbow-colorful-colors-splash-bright-painting-fo.html
Alert	Picture 3	Man focusing	NAPS	People_166_h	N/A
Calm	Picture 1	Blue sky with clouds	Google	N/A	https://freerangestock.com/photos/153325/a-blue-sky-with-clouds.html
Calm	Picture 2	Woman meditating	Google	N/A	https://stocksnap.io/photo/woman-meditating-EWIPFTTJIW
Calm	**Picture 3**	**Monk meditating**	**Google**	**N/A**	https://pixnio.com/people/male-men/man-buddhism-meditation-monk-religion-worship
Excited	**Picture 1**	**Concert**	**Google**	**N/A**	https://www.pickpik.com/crowd-people-music-festival-concert-party-76062
Excited	Picture 2	Rollercoaster	Google	N/A	https://www.flickr.com/photos/masstravel/9127116099
Excited	Picture 3	Diver jumping into swimming pool	NAPS	People_196_h	N/A

Discrete Emotion	Picture Number	Content of Picture	Source	Picture Number from Source	Link (if any)
Relaxed	**Picture 1**	**Girl reading**	**Google**	**N/A**	https://pixnio.com/media/book-reading-relaxation-spring-time-woman
Relaxed	Picture 2	People in hammock	Google	N/A	https://picryl.com/media/hammock-relax-beach-travel-vacation-d77e76
Relaxed	Picture 3	Empty beach sunset	Google	N/A	https://www.pickpik.com/faro-portugal-algarve-most-beach-atlantic-coast-empty-beach-57993
Contentment	Picture 1	Two cats asleep	NAPS	Animals_166_v	N/A
Contentment	**Picture 2**	**Mother and sleeping baby**	**NAPS**	**Faces_079_h**	N/A
Contentment	Picture 3	Tulips in garden	Google	N/A	https://www.wallpaperflare.com/multicolored-flower-garden-with-gray-concrete-pathway-tulips-wallpaper-sjhv
Boredom	Picture 1	Keyring	IAPS	7059	N/A
Boredom	**Picture 2**	**Wooden chair**	**IAPS**	**7235**	N/A
Boredom	Picture 3	Keyboard	NAPS	Objects_245_h	N/A

Study 2b: Picture Descriptions

1. Imagine a picture of a group of men violently attacking a defenseless victim curled up on the ground while onlookers stand nearby (Anger)
2. Imagine a picture of a frail, elderly man lying motionless in bed, his hollow cheeks and sunken eyes reflecting a long, painful struggle with illness (Sadness)
3. Imagine a picture of a snarling wolf baring its sharp fangs, eyes locked in an intense gaze, poised and ready to attack (Fear)
4. Imagine a picture of a cheerful, laughing baby sitting outdoors on a blanket in the grass, clapping her tiny hands (Happiness)
5. Imagine a picture of a dirty toilet bowl containing filthy, brown feces, splattered stains, and soiled toilet paper (Disgust)
6. Imagine a picture of a track athlete leaning heavily over a hurdle, his head resting on his folded arms, body slumped, and posture revealing the toll of exertion after intense effort (Fatigue)
7. Imagine a picture of a vibrant, bustling city square, illuminated by bright billboards and towering skyscrapers, buzzing with energy as people hurry past and cars speed by (Alert)
8. Imagine a picture of a monk sitting on a mat, meditating with his eyes gently closed, bathed in soft, golden light filtering through a canopy and surrounded by lush greenery (Calm)
9. Imagine a picture of a crowd of people at a concert waving their hands energetically as colorful confetti fills the air, capturing the thrill of the moment (Excited)
10. Imagine a picture of a woman lying lazily and comfortably in long, green grass, on her back with knees bent, leisurely reading her book (Relaxed)

11. Imagine a picture of a mother lying smiling beside her sleeping newborn, her arm wrapped gently around his tiny form, their faces close as they share a moment of tenderness (Contentment)
12. Imagine a picture of a single wooden chair standing against a light blue wall in a room with tan vinyl flooring (Boredom)

Study 3

Table S6

Sample demographics for Study 3

	M	*SD*	%
Mean age in years (*SD*)	39.36	12.25	
Gender			
Male			55.72%
Female			43.78%
Non-binary			0.50%
Ethnicity/Race			
White			61.19%
African American			18.91%
Asian/Pacific Islander			7.96%
Native American			1.49%
Latino/a/x			4.48%
Multi-racial			5.47%
Other			-
Decline			0.50%
Education Level			
Some high school			0.50%
High school (diploma, GED, or alternative credential)			15.42%
Some college credit, but less than 1 year of college credit			9.45%
1 or more years of college credit, no degree			13.43%
Associate’s degree (for example: AA, AS)			10.45%

Bachelor's degree (for example: BA, BS)	34.83%
Master's degree (for example: MA, MS, MEng, MEd, MSW, MBA)	13.43%
Professional degree beyond a bachelor's degree (for example: MD, DDS, DVM, LLB, JD)	1.49%
Doctorate degree (for example: PhD, EdD)	1%
Political Ideology	
Very liberal	12.44%
Liberal	24.88%
Somewhat liberal	11.94%
Moderate	20.90%
Somewhat Conservative	12.94%
Conservative	9.95%
Very Conservative	6.97%
Color Blindness	
Yes	1%
No	99%
Not sure	-

Table S7

Results from linear mixed effects regression on absolute affective forecasting error scores, including fixed effects for measurement type (deviation-coded), and random intercepts for participants and items for Study 3

	Absolute affective forecasting error scores		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	0.82	0.72 – 0.92	<0.001***
Dimensional-Discrete	-0.02	-0.13 – 0.09	0.717
Random Effects			
σ^2	0.88		
τ_{00} PROLIFIC_PID	0.08		
τ_{00} items	0.02		
ICC	0.11		
N PROLIFIC_PID	192		
N items	12		
Observations	2303		
Marginal R^2 / Conditional R^2	0.000 / 0.107		

Note. *p < .05; **p <.01; ***p < .001

Table S8

Results from linear mixed effects regression on responses to affect ratings, including fixed effects for type of affect rating (deviation-coded), measurement type (deviation-coded), baseline affect (centered), and random intercepts for participants and items for Study 3

	Affect rating		
Predictors	*Estimates*	*96% CI*	*p*
Intercept	4.82	4.22 – 5.42	<0.001***
Predicted-Actual	0.11	0.03 – 0.18	0.008**
Dimensional-Discrete	-0.14	-0.41 – 0.12	0.289
Baseline affect centered	0.02	0.01 – 0.04	<0.001***
Predicted-Actual × Dimensional-Discrete	-0.02	-0.17 – 0.14	0.847
Predicted-Actual × baseline affect centered	-0.00	-0.01 – 0.00	0.590
Dimensional-Discrete × Baseline affect centered	0.01	-0.01 – 0.03	0.304
Predicted-Actual × Dimensional-Discrete × Baseline affect centered	-0.00	-0.02 – 0.01	0.525
Random Effects			
σ^2	1.84		
τ_{00} PROLIFIC_PID	0.81		
τ_{00} items	1.07		
ICC	0.51		
N PROLIFIC_PID	192		
N items	12		
Observations	4607		
Marginal R^2 / Conditional R^2	0.028 / 0.520		

Note. *p < .05; **p <.01; ***p < .001

Table S9

Results from linear mixed effects regression on responses to affect ratings, including fixed effects for type of affect rating (deviation-coded), measurement type (deviation-coded), age (centered), and random intercepts for participants and items for Study 3

	Affect rating		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	4.83	4.22 – 5.43	<0.001***
Predicted-Actual	0.10	0.02 – 0.18	0.012*
Dimensional-Discrete	-0.09	-0.38 – 0.19	0.523
Age centered	0.01	-0.01 – 0.02	0.258
Predicted-Actual × Dimensional-Discrete	-0.01	-0.16 – 0.15	0.935
Predicted-Actual × Age centered	0.00	-0.00 – 0.01	0.377
Dimensional-Discrete × Age centered	0.01	-0.02 – 0.03	0.552
Predicted-Actual × Dimensional-Discrete × Age centered	-0.00	-0.02 – 0.01	0.572
Random Effects			
σ^2	1.84		
τ_{00} PROLIFIC_PID	0.91		
τ_{00} items	1.07		
ICC	0.52		
N PROLIFIC_PID	192		
N items	12		
Observations	4607		
Marginal R^2 / Conditional R^2	0.004 / 0.520		

Note. $*p < .05$; $**p < .01$; $***p < .001$

Table S10

Results from linear mixed effects regression on responses to affect ratings, including fixed effects for type of affect rating (deviation-coded), measurement type (deviation-coded), number of Sudoku clicks, and random intercepts for participants and items for Study 3

	Affect rating		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	4.90	4.25 – 5.54	<0.001***
Predicted-Actual	0.20	0.05 – 0.35	0.011*
Dimensional-Discrete	-0.22	-0.77 – 0.33	0.435
Sudoku Click Count	-0.00	-0.01 – 0.00	0.528
Predicted-Actual × Dimensional-Discrete	-0.14	-0.45 – 0.16	0.353
Predicted-Actual × Sudoku Click Count	-0.00	-0.01 – 0.00	0.163
Dimensional-Discrete × Sudoku Click Count	0.00	-0.01 – 0.02	0.676
Predicted-Actual × Dimensional-Discrete × Sudoku Click Count	0.00	-0.00 – 0.01	0.349
Random Effects			
σ^2	1.84		
τ_{00} PROLIFIC_PID	0.91		
τ_{00} items	1.07		
ICC	0.52		
N PROLIFIC_PID	192		
N items	12		
Observations	4607		
Marginal R^2 / Conditional R^2	0.003 / 0.520		

Note. $*p < .05$; $**p < .01$; $***p < .001$

Table S11

Results from linear mixed effects regression on responses to affect ratings, including fixed effects for type of affect rating (predicted, actual), measurement type (dimensional, discrete), stimulus valence (pleasant, unpleasant), and stimulus arousal (low, high) (all deviation-coded), and random intercepts for participants and items for Study 3

	Affect rating		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	4.82	4.25 – 5.39	<0.001***
Predicted-Actual	0.10	0.03 – 0.18	0.007**
Dimensional-Discrete	-0.12	-0.40 – 0.16	0.416
Stimulus valence: Pleasant-Unpleasant	0.35	-0.75 – 1.46	0.534
Stimulus arousal: Low-High	-0.63	-1.73 – 0.48	0.264
Predicted-Actual × Dimensional-Discrete	-0.02	-0.17 – 0.14	0.824
Predicted-Actual × Stimulus valence: Pleasant-Unpleasant	0.09	-0.06 – 0.24	0.246
Dimensional-Discrete × Stimulus valence: Pleasant-Unpleasant	0.74	0.59 – 0.89	<0.001***
Predicted-Actual × Stimulus arousal: Low-High	-0.10	-0.25 – 0.05	0.202
Dimensional-Discrete × Stimulus arousal: Low-High	-0.55	-0.70 – -0.39	<0.001***
Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	1.86	-0.35 – 4.07	0.099
Predicted-Actual × Dimensional-Discrete	-0.14	-0.45 – 0.16	0.357

× Stimulus valence: Pleasant-Unpleasant			
Predicted-Actual × Dimensional-Discrete × Stimulus arousal: Low-High	0.18	-0.12 – 0.49	0.243
Predicted-Actual × Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	0.47	0.16 – 0.78	0.003**
Dimensional-Discrete × Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	0.98	0.67 – 1.28	<0.001***
Predicted-Actual × Dimensional-Discrete × Stimulus valence: Pleasant-Unpleasant × Stimulus arousal: Low-High	-0.45	-1.06 – 0.17	0.154
Random Effects			
σ^2		1.76	
τ_{00} PROLIFIC_PID		0.91	
τ_{00} items		0.95	
ICC		0.51	
N PROLIFIC_PID		192	
N items		12	
Observations		4607	
Marginal R^2 / Conditional R^2		0.107 / 0.564	

Note. *p < .05; **p <.01; ***p < .001

Table S12

Results from linear mixed effects regression on responses to affect ratings excluding responses to pictures and picture descriptions I was less confident about. The model includes fixed effects for type of affect rating (predicted, actual), measurement type (dimensional, discrete) and random intercepts for participants and items for Study 3

	Affect rating		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	4.98	4.22 – 5.75	<0.001***
Predicted-Actual	0.08	-0.01 – 0.18	0.078
Dimensional-Discrete	0.06	-0.24 – 0.36	0.696
Predicted-Actual × Dimensional-Discrete	-0.10	-0.28 – 0.08	0.289
Random Effects			
σ^2		1.70	
τ_{00} PROLIFIC_PID		1.00	
τ_{00} items		1.17	
ICC		0.56	
N PROLIFIC_PID		192	
N items		8	
Observations		3071	
Marginal R^2 / Conditional R^2		0.001 / 0.560	

Note. *p < .05; **p <.01; ***p < .001