INTENSE EMOTION REACTIONS PREDICT ENHANCED WELL-BEING AND ADAPTIVE CHOICES

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Robert John Klein

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Psychology

July 2020

Fargo, North Dakota

North Dakota State University
Graduate School

Title

INTENSE EMOTION REACTIONS PREDICT ENHANCED WELL-BEING AND ADAPTIVE CHOICES

By

Robert John Klein

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Michael D. Robinson
Chair

Clayton Hilmert

Mark McCourt

Joel Hecktner

Approved:

July 17, 2020
Date

Mark Nawrot
Department Chair

ABSTRACT

Existing evidence has linked individual differences in emotion reaction intensity to both enhanced and decreased psychological well-being. We propose that this contradiction is related to methodological shortcomings in some existing research. We present a novel emotion reactivity task capable of addressing these methodological shortcomings by continuously measuring the subjective intensity of individual emotion episodes with high temporal resolution.

Four studies were conducted (total n = 499). In Studies 1, 2, and 4, participants continuously reported their emotions while viewing objectively pleasant or unpleasant images. Thousands of reaction intensities were coded using algorithms developed for this purpose. We expected that people showing more intense emotion reactions, regardless of valence, would report greater subjective well-being in the lab and in daily life. One reason that such situationally-congruent reactions might be beneficial is that that they enable more flexible situationally-appropriate behavior. In Study 3, participants were asked to rate their emotional responses to pleasant and unpleasant images. Following this, people choose a location for their Self avatar within a computerized environment that included one image of each valence. We expected that the tendency to report intense emotion responses to these images would predict both adaptive location choice and subjective well-being.

Results confirmed most major hypotheses: more intense reactions to both positive and negative stimuli were predictive of greater subjective well-being in the lab and in daily life, and analogous reactivity patterns were associated with more flexible, adaptive avatar placement. The results suggest that a key feature of maladaptive emotion generation systems (and lower well-being) may not be overly intense reactions as has been suggested, but a failure to flexibly adapt emotion output to match changing circumstances.

ACKNOWLEDGMENTS

First and foremost, I would like to thank Dr. Michael Robinson for his guidance, knowledge, patience, creativity, and foresight throughout my dissertation process (more accurately, throughout my entire training at NDSU). I also would like to acknowledge the research assistants that helped collect the present data. I thank Ganesh Padmanabhan for assisting me with E-Prime programming. I am pleased do acknowledge Graham Anderson, one of my oldest friends, for his Python programming and support with encoding the scoring algorithms. All these people were essential to the success of this project. I would also like to thank my committee for their time and support.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Emotional experiences are a pivotal component of psychological health. Indeed, emotions often define psychological health. For instance, a person's subjective well-being, or "happiness", is traditionally defined as a combination of typical emotional experiences and positive judgments about one's life as a whole (Diener, 2000). Given this emotion/well-being link, there is good reason to think that individual differences in emotional reactivity should play a key role in psychological well-being. Consistent with this point, significant evidence suggests that "healthy" emotion reactions (e.g., ones associated with higher well-being and less psychopathology) are characterized by less intense peak emotion (e.g., Gruber, Kogan, Quoidbach, & Mauss, 2013). Relatedly, there is a sizeable body of evidence connecting lower well-being with more intense emotion reactions to unpleasant events or circumstances (Gross, Sutton, & Ketelaar, 1998; Zelenski & Larsen, 1999).

Problematically, however, there is growing evidence suggesting that typical measures of emotion reactivity – including the measures that form the basis of the majority of research investigating emotion reactivity and well-being – may provide a somewhat misleading picture of the emotion that occurred. This is so because most typical emotion reaction measures fail to model emotion reactions as *dynamic* experiences that change over time (Davidson, 2015; Frijda, 2007). For example, in traditional emotion reaction assessments, a subject might be asked to summarize their peak state affect after some emotion induction. In contrast to this monolithic and static perspective, affective science has begun to view emotion reactions as multidimensional processes that unfold over time (Davidson, 2015; Frijda 2007; Houben, van Den Noortgate, & Kuppens, 2015; Scherer, 2000). From this dynamic perspective, when a person summarizes the size or strength of some reaction, this report could reflect an amalgamation of peak reaction intensity along with other factors such as how long the emotion reaction lasted (Davidson, 1998). Such a confound is a significant problem when reaction intensity and reaction duration are viewed as very different phenomena with very different causes and downstream impacts on psychological health.

Given these and other limitations, it is not surprising that recent emotion reactivity research has begun using more precise dynamic methods, and these studies have reported results that are not aligned with traditional views of emotion reactivity and its relationship to psychological health. Specifically, some recent work has suggested that more intense emotion reactions (even to negative stimuli) may actually be a sign of *enhanced* mental health (Höflich, Michenthaler, Kasper, & Lanzenberger, 2018; Schaefer et al., 2013; Taubitz, Robinson, & Larson, 2013; Waugh, Thompson, & Gotlib, 2011). Given such contradicting evidence (discussed in greater detail below), I propose that we need to develop new and creative methods for assessing individual differences in emotion reaction intensity.

Along these lines, there is growing awareness among affective scientists that we need to model emotion reactions as dynamic constructs if we are to understand the true relationship between emotion reactivity and well-being (Kuppens & Verduyn, 2017). Researchers attempting to address this gap have suggested that emotion reactions are actually comprised of multiple distinct features such as reaction latency, peak reaction intensity, rise-time, reaction duration, etc. (Davidson 1998; Kuppens & Verduyn, 2017). As we begin to view emotion reactions as dynamic, multifaceted entities, key new questions have arisen: what exactly does a mentally healthy emotion reaction look like over time? What are the well-being impacts of individual differences in dynamic emotion features such as peak reaction intensity or peak reaction duration? Might these impacts vary by stimulus valence? Do people who tend to have more intense emotion reactions to positive stimuli also have more intense reactions to negative stimuli? Answers to these questions could provide key insights into how we might effectively target emotion reactivity to improve mental well-being (Kuppens & Verduyn, 2017).

This newer research on “affective chronometry” (Davidson, 2000) or “emotion dynamics” (Houben et al., 2015) has uncovered some consistent findings. For instance, individual differences in the duration of one’s typical emotion reactions appear to consistently predict psychological well-being, such that individuals with longer reactions tend to exhibit lesser well-being (Davidson, 2000; Fredrickson & Jointer, 2002; Houban et al., 2015; Lyubormirsky, King, & Diener, 2005). Moreover, given that well-being is largely defined in terms of emotions (Diener & Biswas-Diener, 2011), it should not be surprising that several research teams have suggested causal links between affective dynamics and subsequent well-being (Kuppens & Verduyn, 2017; Van de Leemput et al., 2014; Wichers & Groot, 2016).

Peak reaction intensity is among the most commonly explored individual difference parameters in affective dynamics (Houben et al., 2018). Research using diverse methods (and accompanying theory) has consistently linked patterns of reaction intensity to mental well-being (Kuppens & Verduyn, 2018). Although this research has often produced opposing results (as reviewed below), a recent meta-analysis concluded that more extreme emotion reactions, whether pleasant or unpleasant, are linked to lower levels of well-being (Houben et al., 2015). I will propose that this specific model has several key flaws related to temporal resolution and features of the activating stimulus. I will also show that methods using high temporal resolution have often demonstrated the opposite pattern.

Resolving the emotion intensity/well-being question matters because if it were to become evident that intense emotion reactions are not a marker of lowered well-being, and that longer duration emotion reactions are actually the key predictor of lowered well-being, this realization would suggest specific changes in interventions designed to improve emotion regulation and related well-being. For instance, maladaptive regulation approaches specifically associated with elongated emotion reactions could be particularly important targets for intervention. The present research program begins to address these questions. In this connection, I explore the relationship between individual differences in dynamic emotion reaction intensities and subjective well-being using a novel approach to quantifying subjective emotion reactions over time with high temporal resolution.

In the next section, I briefly review the opposing evidence and theory concerning the relationship between dynamic affective reactivity and psychological markers of well-being. Following this review, I describe the present program of research which is designed to address specific gaps in this literature.

Subjective Well-being and Reaction Intensity

As I have said, there is evidence suggesting that "healthy" emotion reactions (e.g., ones associated with higher well-being and less psychopathology) involve less intense emotion reactions (Gruber et al., 2013; Houben et al., 2015). In terms of *lower* well-being and higher emotion reactivity, for example, there is a sizeable body of evidence connecting trait neuroticism with increased reactivity to aversive experiences (Gross, Sutton, & Ketelaar, 1998; Zelenski & Larsen, 1999). These links have also been supported in studies specifically modeling questions related to emotion dynamics (Heller, Fox, & Davidson, 2018; Houben et al., 2015). These negative links between emotional intensity and well-being are sometimes stronger with respect to negative emotions. Yet, similar results have been reported for positive emotions as well. For instance, a recent meta-analysis reported that higher well-being individuals often display lesser intensity deviations from baseline with respect to both their positive and negative emotional reactions (Houben et al., 2015).

Several theoretical orientations could explain this negative association between emotion intensity and well-being. In terms of overreactions to negative stimuli, Wichers, Wigman, and Myin-Germeys (2015) suggest that experiencing more intense deviations from baseline feelings is problematic for well-being because it contributes to chronic overreactions to stressors, which can add up over time. Such overreactions to aversive stimuli have also been identified as key risk factors for future psychopathology (Charles, Piazza, Mogle, Sliwinski, & Almeida, 2013). Given that emotional states have been shown to bias the interpretation of subsequent events (Kanske & Kotz, 2012), these over-reactions may build upon themselves to perpetuate and exaggerate negative emotional states in a sort of downward spiral (Wichers et al., 2015).

In terms of the intensity of positive emotion reactions, the data also seem to show a link between higher well-being and *less* reactivity to positive events (Houben et al., 2015; Oishi, Diener, Choi, Kim-Prieto, & Choi, 2007; Rueschkamp et al., 2018). However, the supporting theory is relatively sparse. One possible explanation is that people with higher well-being may become habituated to positive events, and so tend to have a less extreme momentary reaction to them (Oishi et al., 2007; Rueschkamp et al., 2018).

As I have alluded to above, there is also a body of evidence, albeit less extensive, suggesting the opposite relationship between well-being and markers of emotion reaction intensity. For example, higher well-being has been associated with more pronounced eye blink startle responses following positive and negative stimuli (Schaefer et al., 2013), and also more extreme subjective reactions to both positive and negative images (Waugh et al., 2011). In addition, higher well-being (in the form of extraversion) has been repeatedly linked to more intense reactions to positive stimuli (e.g., Larsen & Ketelaar, 1991). Also, greater variability in cortisol levels, a marker of negative emotion reactivity, has been characteristic of enhanced psychological adjustment in general (Mikolajczak et al., 2010). Similarly, *blunted* peak cortisol responses to stressors have been implicated in chronic stress and lower well-being (Sapolsky, 2004). Relatedly, depression has been a reliable predictor of less intense reactions to both pleasant and unpleasant events (Höflich et al., 2018; Taubitz et al., 2013).

One key argument supporting a positive association between well-being and reaction intensity is that robust emotion reactions generally serve some adaptive purpose (Nesse & Ellsworth, 2009). For instance, emotions often motivate individuals to address and/or cope with important threats or rewards (Frijda, 2004). People who tend to have robust “context-dependent” emotion reactions seem better able to cope with difficulties and behave adaptively (Cheng, 2001; Waugh et al., 2011). Such context dependent reactions are a key indicator of Psychological Flexibility. Psychological Flexibility is an individual difference involving adaptive tendencies such as the ability to respond to situational demands, flexibly shift mindsets or behavior, and a willingness to experience reality as it is (Kashdan & Rottenberg, 2010). As an individual difference, psychological flexibility seems to be a key determinant of psychological well-being and mental health (Kashdan & Rottenberg, 2010).

Relatedly, blunted emotion reactions may reflect attempts to suppress or escape emotional experiences. Such avoidant tendencies may lead to lowered well-being given that they have been shown to be ineffective in the long run (Gross & John, 2003). Moreover, avoidant tendencies have been shown to significantly increase the overall quantity of negative emotions, possibly by increasing vigilance for threats and disappointments (Hayes, Strosahl, & Wilson, 2012; Tull, & Roemer, 2007).

In sum, there seems to be substantial empirical and theoretical support for both positive and negative links between reaction intensity and well-being. This basic confusion about the characteristics of a healthy emotion reaction remains a problem for clinical researchers and therapists attempting to develop solutions for enhancing emotional well-being. What should not be in question, however, is that existing approaches to studying emotion intensity have limitations and it is to these limitations that I now turn.

Emotion dynamics studies reporting negative associations between individual differences in reaction intensity and well-being have often used retrospective reports measuring a sort of average emotion intensity felt over the course of an entire day. These studies are not measuring discrete emotion reactions in their entirety, and thus are making assumptions when differentiating intense emotion reactions from other dynamic reactivity components. This is true because retrospective summations of emotion reactions may not differentiate between a high intensity emotion reaction and other features of the reaction such its duration (Davidson, 1998). Similar daily designs have also been used in EMA studies of state affect. This non-retrospective approach is more desirable, but is also somewhat flawed in that it makes assumptions about the timing or intensity of eliciting stimuli and it is uncertain whether any objectively provocative stimulus was present at all (Houben et al., 2015; Koval et al., 2016). Furthermore, such EMA and daily affect summary studies typically operationalize reaction intensity using standard deviations or closely related metrics calculated on repeated affect assessments. Given the lack of control over activating event intensity and timing, these measures may be capturing tendencies to overreact to minor events (Trull et al., 2015), or even sustain an emotion reaction over time. Similarly, higher standard deviations could simply occur because a person experiences a greater frequency of stressors. I discuss these limitations in greater detail in the General Discussion.

Research designs showing positive associations between well-being and greater reactivity are also limited in several ways. First, these studies are often focused on physiological measures or other indirect approximations of subjective emotion experiences. This matters because it is principally the subjective component of emotion experiences that links emotion to human quality of life. We simply need to do a better job of assessing subjective emotional experiences with high temporal resolution and with controlled stimuli. Second, this reaction intensity research is also largely focused on ill-being measures like depression and anhedonia rather than well-being measures such as flourishing or life satisfaction. In sum, extant research has provided important clues about the relationship between the extremity of emotion reactions and well-being, but many important questions remain to be addressed. The present research builds on prior work by developing and testing methods capable of addressing some of these gaps.

The Present Studies

We completed four studies. Our intention was to assess the relationship between emotion reaction intensity and subjective well-being in paradigms designed to assess the dynamic components of emotional experience and reactivity.

Study 1

Study 1 deployed a novel emotion reactivity paradigm called the Dynamic Affective Reactivity Task (DART). The DART was designed to assess the peak emotion intensities of discrete subjective reactions to controlled stimuli. Study 1 also administered several well-being questionnaires. Psychological Flexibility theories suggest that mental well-being should be positively related to the intensity of one's reactions to provoking stimuli (Kashdan & Rottenberg, 2010; Waugh et al., 2011). We expect this to be true in Study 1 as well. We also expect that such relations should be present with respect to emotion reactions of both valences (Waugh et al., 2011).

Study 2

We designed Study 2 as a close, but not identical, replication of Study 1. Such a replication is desirable given the novel nature of the DART and the hypotheses. This is also important because it is not typical to suggest that the intensity of emotion reactions (including negative reactions) should be positively correlated with well-being (see, for instance, the meta-analysis by Houben & Kuppens, 2015).

Study 3

This study addressed a potential mechanism that could help explain the results of Study 1 and 2. Study 3 was theoretically rooted in the idea that emotion reactions would not exist if they were not useful (Nesse & Ellsworth, 2009). Because emotions likely solve some problem (Nesse & Ellsworth, 2009), people who show larger emotion reactions to objectively positive and negative stimuli may be more capable of solving problems than people with more blunted emotion reactivity systems. One fundamental problem that all organisms encounter is that their environments are constantly changing (Nesse & Ellsworth, 2009)—sometimes a threat is present, but at other times a reward is present. To survive in such a drastically heterogeneous environment, humans must both approach rewards while also avoiding threats. To accomplish such situation-specific behavior, people must also possess the requisite motivations (Gollwitzer & Bargh, 1996). A primary source of this approach/avoidance motivation comes from subjective emotional experiences (Carver & White, 1994). This is true because people are generally motivated to experience pleasant feelings and to avoid experiencing unpleasant feelings (Miller, 1944). To the extent that these feeling-based systems are working, the person should be better able to make adaptive choices in how they behave in response to the situation.

When we think in this way, a potential mechanism by which greater emotion reactivity should lead to greater personal well-being (as expected in Studies 1 and 2) becomes apparent. Specifically, if emotions exist because they motivate adaptive behavior, then having more intense emotional experiences should lead to more adaptive behavior in a changing environment. And, this matters for the present hypotheses because flexible or adaptive behavior should be linked to personal well-being (Kashdan & Rottenberg, 2010). Such adaptive behavior should include, for instance, successfully avoiding negative or problematic stimuli and successfully approaching rewarding ones (Nesse & Ellsworth, 2009).

To test for the presence of a behavioral-flexibility mechanism connecting emotion reactivity and well-being, we designed a paradigm that principally involved a changing environment and recorded people's emotional and behavioral responses to that environment. To test whether people were effectively motivated to approach positive stimuli and avoid negative stimuli (in a somewhat implicit way), we first assessed the extent to which people experienced each image as being pleasant or unpleasant. Then, we asked people to freely place an avatar representing "the self" in the space between the images. We expected that people who showed more intense emotion reactions to the normed stimuli should also place themselves in more adaptive/situationally-congruent ways (defined as closer to the stimulus they labeled as pleasant and further from the stimulus they labeled as unpleasant). More importantly, we expected that this behavioral-flexibility would be associated with higher reports of subjective well-being.

Study 4

All studies to this point have assessed well-being using laboratory-based measures. To enhance the validity of the present findings, Study 4 adopted a daily diary perspective on DART performance. The key question was whether higher intensity reactions to both valences (collapsing across valence) would be predictive of greater daily well-being, even up to 4 weeks later.

STUDY 1

Method

Participants and General Procedures

Power Considerations. The present within-subject design was chosen precisely because it is a powerful one (Loersch & Payne, 2016). Also, the multilevel modeling strategy used here should provide the best possible estimates of the key parameters (Nezlek, 2012). Importantly, the precise calculation of power for multilevel models is complex as well as error-prone (Nezlek, 2012). We therefore based our decisions about sample size on recommendations in the literature (Nezlek, 2012; Scherbaum & Ferreter, 2009), which suggest that sample sizes of 100 participants with at least 10 observations per person should provide good power for our multilevel hypotheses (Ohly, Sonnentag, Niessen, & Zapf, 2010). To obtain sample sizes of 100 or more people, we conducted the laboratory portion of each study for two weeks, as procedures of this type have resulted in sample sizes of 100 or more in previous studies within the lab.

Participants (n = 134, M age = 18.99 SD = 1.74, 73.13% women, 88.81% Caucasian) were undergraduates at North Dakota State University who were recruited using SONA management software and compensated with course credit. Before commencing, all studies were approved by the Institutional Review Board at North Dakota State University, and participants were informed of their ability to terminate the experiment at any time for any reason. These students arrived at the laboratory in groups of six or fewer and completed an affect dynamics task using E-Prime 2.0 as well as a subjective well-being assessment using MediaLab software. All computers are equipped with 12.5” x 16.5” (1280 x 1024) monitors of an identical model (LG Flatron ME 20CR-BF).

Subjective Well-Being (SWB) Assessment

See Appendix A for a full description of all measures. We assessed subjective well-being levels using the Flourishing scale and the positive component of the Scale of Positive Emotion and Negative Emotion. These scales were selected because together they capture psychosocial prosperity across a theoretically broad range of affective, experiential, social, and eudemonic aspects (Diener et al., 2010). The Flourishing scale (FS) has been a reliable instrument and has correlated highly with established global well-being measures (e.g., satisfaction with life). The scale incorporates basic need satisfaction in terms of personal meaning and social support, as well as prosociality, optimism, engagement with one's life, and self-esteem. On the FS (M = 5.72, SD = .72, alpha = .87), participants rated the extent to which they agree (1 = strongly disagree; 7 = strongly agree) with 8 statements related to general success in key life domains such as "I lead a purposeful and meaningful life" and "my social relationships are supporting and rewarding". For both scales, item level scores were averaged to create a single score.

The Positive Emotionality (PE) component of the Scale of Positive and Negative Emotion (Diener et al., 2010) is a commonly employed well-being instrument that asks participants to indicate how often (1 = very rarely or never; 5 = very often or always) they have experienced each of six broad, non-specific positive feeling states (Positive, Happy, Contented, Pleasant, Good, Joyful) over the past four weeks (M = 3.86, SD = .50, alpha = .79). We focus here on well-being as defined by the presence of positive indicators (PE) rather than the absence of negative indicators. This was done to draw a sharp distinction between well-being and psychopathology, the latter being principally defined by the presence of negative emotional symptoms.

Because we were principally interested in the relationship between emotion reactivity and overall subjective well-being, we created a composite well-being score by standardizing and averaging the PE and FS scales (M = -.004, SD = .87). The FS and PE components were correlated at r = .58.

Dynamic Affective Reactivity Task (DART)

Procedure Synopsis. To meaningfully contribute to affective science, the present paradigm needed to capture subjective emotion responses to controlled/normed stimuli with temporal resolution high enough to determine the precise time-point and emotion rating at which people's reactions peaked. To meet these requirements, we relied on a modified version of a continuous tracking procedure previously developed in our lab (Klein & Robinson, 2019; Klein, Towers, & Robinson, in press). Participants viewed 10 pleasant and 10 unpleasant images selected from the Nencki Affective Picture System (NAPS: Marchewka, Zurawski, Jednorog, & Grabowska, 2014) that appeared in the center of their computer screens. These images were presented for 5 seconds after which they were replaced by a black rectangle of the same dimensions. Instructions presented prior to task were as follows:

> During this section of the experiment (about 15 minutes), we would like you to report your EMOTIONAL REACTIONS TO DIFFERENT IMAGES. We are interested in your basic feelings, that is how pleasant you feel or unpleasant you feel, in response to the image.
>
> To rate your feelings, you will use the rating scale that will appear to the right on your screen and the KEYBOARD. Specifically, you will use the "UP" and "DOWN" arrow buttons. When you notice your feelings changing in an

unpleasant direction, push the [NEGATIVE] arrow key, and when you notice your feelings changing in a pleasant direction, push the [POSITIVE] arrow key.

So, every time your feelings change, you should press one of the arrow keys until the rating bar reflects how you feel.

On this emotion rating task, simply try your best to rate the changes in your feelings, and do this EVEN AFTER THE SLIDES HAVE DISAPPEARED. That is, after the image disappears, continue to monitor and rate your current feeling state for the FULL DURATION OF THE TRIAL.

Monitoring and rating your feelings is a strange experience at first. However, human emotions are ALWAYS CHANGING, so try to notice and report these changes.

Remember, whether you are currently seeing an image or not, you should always be monitoring and rating your changing feelings using the arrow keys to change the rating bar.

After each image rating period, you will be asked to make a summary judgment about your feelings regarding the image you just saw. And then you will start at your feeling baseline for the next trial.

In addition to these pre-task instructions, continuously on-screen instructions asked participants to “please monitor your feelings and rate them as they change or stay the same”. While images were displayed, participants rated their feelings using the keyboard arrow keys and a sliding visual rating bar located in a right-justified vertically-oriented rectangle (see Figure 1 for a grayscale version of the task environment). The top emotion label (Very Pleasant or Very Unpleasant) was counterbalanced across participants. This rating procedure was conceptually based on a previously validated continuous rating system (Levenson & Gottman, 1983; see also Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). Rating bar locations were recorded every 100 milliseconds and ranged from -500 to +500, depending on the participants' positioning of the bar. Participants received semi-continuous feedback concerning their current bar location.

Figure 1. DART environment. Actual environment was in color. Image is for descriptive purposes only- not actual NAPS stimuli.

Affective Stimuli. Ten images of each valence were selected based on NAPS norm data for valence (1 = very negative, 9 = very positive) and arousal (1 = relaxed, 9 = aroused). Images were matched for arousal and valence extremity. We first confirmed that the pleasant stimuli we selected (M = 7.07) were indeed more pleasant than the unpleasant images (M = 3.06), $F(1, 18) = 593.60$, $p < .001$, PES = .97. To match positive and negative images for valence extremity, we created extremity scores by calculating the absolute value of image valence norms minus 5 (M pleasant = 2.07, M unpleasant = 1.94). Positive and negative image sets did not significantly differ with respect to their scores for valence extremity, $F(1, 18) = .64$, $p = .435$, PES = .03, or arousal, $F(1, 18) = .27$, $p = .610$, PES = .01. The content of these images varied and included faces, people, objects, animals, and landscapes.

DART Trial Procedure. Participants were first randomly assigned to either the "top good" condition, where the "Very Pleasant" label was located at the top of the rating bar, or to the "top bad" condition. The task included 20 total trials—10 trails of each valence. Before each image appeared, a "Get Ready!" message was presented for 3 seconds. Affective images were sized 10.5 inches x 9.25 inches and appeared for 5 seconds, after which the images were replaced by a black rectangle for 10 seconds.

Data Manipulation and Cleaning

Reaction Intensity Quantification. As indicated above, rating bar location was assessed ten times per second. Thus, each trial produced a stream of approximately 150 emotion intensities and there were 2680 trials. To evaluate our hypotheses, each trial or stream was coded to determine a peak reaction intensity. Two interlocking algorithms were developed to automate this coding process. These algorithms were capable of objectively coding thousands of discrete emotion reactions that occurred and dissipated at differing times for differing images and participants.

To determine each trials' peak reaction intensity, algorithms (described in detail below) were calibrated to determine the precise times at which a) significant increases in emotion intensity began (termed "React Start Time"), and b) these increases ended (termed "Peak Time"). For each trial, the Peak Time point was associated with a corresponding emotion rating. In terms of our hypotheses, the key outcome of these algorithms was this "Peak Intensity" value, or the rating bar position that corresponded with the peak time algorithm output.

Algorithms. Two algorithms were applied to each stream of emotion rating scores to determine this Peak Intensity value. First, difference scores were generated by subtracting each of the 150 emotion ratings from the subsequent rating. Each difference score was associated with a sample number and specifically with the sample number of the first datum in the change score. Because emotion was sampled every 100 milliseconds starting with stimulus onset, sample numbers give an accurate assessment of the time elapsed following stimulus onset. To allow for direct comparisons between positive and negative emotion ratings, negative peak scores were reverse scored. Thus for both positive and negative images, a positive difference score indicates an emotion stream that is becoming more intense.

Reaction Start Times were defined in terms of the first sample number that was associated with both a) two change scores in a row that were greater than zero, as well as b) three sequential change scores (two of which were included in the initial criterion) whose average was greater than four. This algorithm avoids small changes or inconsistent changes being coded as a reaction start. To be coded as a reaction start, the changes indicated above must be positive, reflecting increases in emotion intensity. By contrast, negative change scores associated with pleasant images were determined to be inadvertent and/or idiosyncratic. Because emotion rating streams were sampled ten times per second, the start time sample number was multiplied by 100 to create a Reaction Start Time in millisecond units. *React Start Emotion* was simply defined as the emotion rating corresponding to this sample number.

Peak Times were conceptualized as the time point at which a given emotion reaction had reached its peak intensity. Algorithmically, peak times were defined as the first sample following Reaction start time associated with both a) three emotion change scores in a row that were either zero or negative in combination with b) the average of the following seven change scores being less than one. Criterion “a” would mean that the reaction was no longer significantly increasing–i.e., the reaction had either flattened or started returning to baseline. The average-of-seven component (criterion b) was included because people occasionally exhibit a brief pause in their reaction (e.g., for 600 ms), but then continue to report significant increases in emotion intensity. *Peak Intensity* was simply defined as the emotion rating corresponding to the peak time sample number.

These algorithms were applied to the raw data using a scoring program written in Python version 2.78 for this purpose. This program read CSVs containing each trial's emotion intensity difference scores as output by E-Prime 2.0. The Python scoring program then generated a CSV file that included each trial's two key time points and the peak intensity rating.

Data Cleaning. To prepare the data for analysis, the first step was to verify and clean the algorithm output. To verify the validity of specific algorithm outputs (e.g., reaction start times), we randomly sampled and visually compared the Python outputs to time/emotion plots generated by Excel. The Python output closely matched the visual coding on 100% of the trials that were sampled.

However, the algorithm was not able to code both a reaction start time and a peak time on 8.5% of trials. Visual inspection of these trials suggested that the vast majority of uncodable trials were the result of subjects simply indicating no coherent reaction to NAPS images. Because emotion reactions are the theoretical focus of the present studies, we dropped these trials from the main analysis.

Results

Preliminary Models

Before addressing key hypotheses, the present findings offered a unique opportunity to address a question of some interest. Specifically, a basic question in emotion dynamics is whether individual differences in emotion reactivity patterns operate similarly or oppositely across valences (Davidson, 2015). For instance, successful emotion regulation efforts could lead to less intense negative emotions and more intense positive emotions (Gross, 2015), and this dynamic would result in a negative correlation across valences. However, psychological flexibility (PF) theory suggests that the opposite might be true–because emotion reactivity to stimuli of both valences may be beneficial (Waugh et al., 2011), people who tend to show more intense reactions to negative stimuli should also show more intense reactions to positive stimuli.

To address this question, we created an OLS dataset by collapsing across DART trials to create, for each subject, an average Peak Intensity score for positive stimuli and an average Peak Intensity score for negative stimuli. We then correlated these two measures with each other and the correlation was definitively positive, $r = .47$, $p < .001$. This result aligns with PF perspectives on emotion reactivity (Waugh et al., 2011). Healthy/flexible emotion systems should generate robust reactions regardless of stimulus valence (Waugh et al., 2011). These results are also consistent with the existence of a common system that might underlie flexible, situation congruent emotion reactivity. For instance, psychological flexibility may be a product of cognitive systems that tend to remain mindful and present with current experience, regardless of whether this experience is pleasant or unpleasant (Kashdan & Rottenberg, 2010). Of course, it is possible that this correlation could be caused by response styles such as a tendency to use scale extremes. However, such response tendencies would not account for the relationship between emotion reactivity and subjective well-being that is documented below.

Hypothesized Effects

Based on my literature review, I have formulated my hypotheses with emotion reactivity preceding or influencing SWB (e.g., Kuppens & Verduyn, 2017). However, the present emotion assessment tools forced us to model emotion reactivity as the dependent variable. Given that regression-based results are causally agnostic, this modeling approach is arbitrary and does not influence our interpretation of the results.

The DART output contained variables that were repeatedly measured within-subjects. Data of this type is best analyzed using Multilevel Modeling (MLM; Nezlek, 2008) procedures. MLM is essentially a typical regression equation that, when calculating error for the key slopes and intercepts, takes into account individual differences in these slopes and intercepts. Through the use of such techniques, more random error is explained and the effects of interest can become more apparent (Nezlek, 2008).

To assess whether larger emotion reaction intensities were linked to higher SWB composite score levels, as well as whether this relationship interacted with image valence, we performed a MLM using SAS Proc Mixed (Singer, 1998). This model included peak intensity (dependent variable), SWB composite scores (z-scored level 2 independent variable), valence (level 1 independent variable; coded -1 for negative images and +1 for positive images), and the valance by SWB interaction term. To directly compare Peak Intensity scores across valence, all Peak Intensity scores associated with negative stimuli were multiplied by -1 (this was appropriate given that a neutral emotion rating was zero, and intensity scores ranged from +500 to -500).

The intercept of this model was significantly non-zero, $b = 263$, $t = 38.8$, $p < .001$, 95% CI = 250 to 277, indicating a moderate to strong average Peak. There was also a significant effect of valence, $b = -53.25$, $t = -12.82$, $p < .001$, 95% CI = -61.27 to -45.22, indicating that positive images were associated with a smaller mean intensity (M positive = 211) than negative images (M negative = 318).

More importantly, this model revealed the hypothesized "level 2 main effect" of the SWB composite on Peak Intensity scores. As expected, higher Peak Intensity scores were associated with higher SWB levels, $b = 15.70$, $t = 2.30$, $p = .023$, 95% CI = 2.17 to 29.23. Estimated means revealed that people reporting higher (+1 *SD*) SWB also experienced greater average Peak Intensities ($M = 279$) than people with lower (-1 *SD*) SWB levels ($M = 248$). Also supporting our hypotheses, image valence did not significantly interact with this effect, $b = -2.47$, $t = -0.59$, $p = .558$, 95% CI = -10.73 to 5.79.

We also re-ran this MLM with each SWB scale individually substituted for the composite score. The model including Positive Emotionality (PE; Level 2 independent variable) revealed a significant effect of valence, $b = -53.45$, $t = -12.75$, $p < .001$, 95% CI = -61.67 to -45.23, a nonsignificant effect of PE, $b = 10.41$, $t = 1.49$, $p = .138$, 95% CI = -3.38 to 24.20, and a nonsignificant valence by PE interaction, $b = 1.80$, $t = 1.42$, $p = .671$, 95% CI = -6.52 to 10.12. The model including Flourishing revealed a significant effect of valence, $b = -53.37$, $t = -12.83$, $p < .001$, 95% CI = -61.53 to -45.21, a significant effect of FS, $b = 18.24$, $t = 2.60$, $p = .011$, 95% CI = 4.35 to 32.13, and a nonsignificant interaction, $b = -6.48$, $t = -1.50$, $p = .134$, 95% CI = -14.97 to 2.00. Together, the Study 1 results suggest that more intense emotion reactions are generally predictive of higher SWB, and that this relationship is present in reactions to both pleasant and unpleasant stimuli.

STUDY 2

Study 2 was designed as a close replication of Study 1, with some minor changes to the DART trial procedures designed to increase the temporal resolution of responding.

Methods

Participants and General Procedures

Power considerations and general procedures for Study 2 were identical to those described in Study 1. Participants (n = 139, M age = 19.03; SD = 1.27, 53.24% women, 87.77% Caucasian) were recruited in the same manner, and people completed the DART before filling out SWB questionnaires.

Subjective Well-Being Assessment

The subjective well-being instruments were identical to those described in Study 1. Subjective well-being levels were measured using the Flourishing Scale (FS; M = 5.66, SD = .97, alpha = .93) and the positive component of the Scale of Positive Emotion and Negative Emotion (PE; M = 3.85, SD = .62, alpha = .85). See Study 1 methods for a more detailed discussion of these measures. We were interested in the relationship between emotion reactivity and overall subjective well-being, and therefore a composite well-being score was created by standardizing and averaging the PE and FS scales (M = .021, SD = .94). The FS and PE components were correlated at r = .69.

Dynamic Affective Reactivity Task

Procedure Synopsis. Again, the DART was designed to capture subjective emotion responses to controlled/normed stimuli while achieving the temporal resolution necessary to isolate the peak intensity in the constant stream of affect. In a randomized order, participants viewed 10 pleasant and 10 unpleasant images. Unlike Study 1, these images were selected from the International Affective Picture System or IAPS (Lang, Bradley, & Cuthbert, 1997). The IAPS slides are commonly used to induce emotion in the lab and have been shown to alter the subjective, behavioral, and physiological components of both positive and negative emotion generation systems (Bradley & Lang, 2007; Lang, 1995; Lang & Bradley, 2010; Mikels et al., 2005). Participants viewed each provocative image for 5 seconds, followed by a neutral IAPS image presented for 10 seconds. Neutral images were included to help create an emotion reaction environment that more closely resembles daily life in that daily life often involves a continuous stream of pleasant, unpleasant, and neutral stimuli (Heller et al., 2018). Participants were required to rate their emotions using a computer mouse instead of the keyboard. Instructions were as follows:

> During this experiment (about 10 minutes), we would like you to pay careful attention to your feelings as you see different images on the screen.
>
> We are interested in your basic feelings, meaning how pleasant or unpleasant you feel during the task. To rate your feelings, you will use a rating scale that will appear to the right on your screen and the COMPUTER MOUSE.
>
> Simply watch the images as they roll by, and do your best to sense and rate your current feelings. Try to simply feel your feelings and rate them, as opposed to rating how you "think" you "should" be feeling! Monitoring and rating your feelings is a strange experience at first. However, you will get better as you go through the task.

The Study 2 DART environment was identical to that used in Study 1 (See Figure 1). Rating bar locations were recorded every 100 milliseconds and ranged from -500 to +500, depending on the participants' positioning of the rating bar.

Affective Stimuli. Image selection procedures were similar to those used in Study 1. Ten images of each valence/pleasantness were first selected. Selections were made using IAPS SAM-based norm data for image pleasure (1 = "very unhappy" experience while viewing, 9 = "very happy" experience while viewing) and image arousal (1 = a sleepy looking manikin, 9 = a manikin with an exploding stomach: Bradley & Lang, 1994). We first confirmed that the pleasant stimuli we selected (M = 7.46) were indeed more pleasant than the unpleasant images (M = 2.46), $F(1, 18) = 1247.36$, $p < .001$, PES = .99. To match positive and negative images for affective extremity, we created extremity scores for each image by calculating the absolute value of the pleasantness norm minus five (M pleasant = 2.46, M unpleasant = 2.43). Two ANOVAs were calculated to ensure that the pleasant and unpleasant image sets did not significantly differ by affective extremity, $F(1, 18) = .05$, $p = .820$, or arousal, $F(1, 18) = .07$, $p = .799$, PES = .00. Pleasurable affective images featured scenes such as sailing or cute animals, and unpleasant images featured disturbing scenes such as physical attacks, car accidents, or crime scene photos.

Each DART image was randomly paired with a neutral image. Thus, 20 slides with neutral IAPS SAM norms were selected so that the emotion reactivity environment was constantly in flux (M pleasantness = 4.93, SD = .27, M arousal = 3.32, SD = .34), rendering the paradigm sensitive to change. Neutral images featured benign or nondescript objects such as a typical coffee mug, a pen, a towel, etc.

DART Trial Procedures. Participants were randomly assigned to either the "top good" condition or the "top bad" condition. The task included 20 trials including 10 images of each valence. Unlike Study 1, the present paradigm did not use a "get ready" image. Rather, E-Prime simply presented affective images for 5 seconds, followed by 15 seconds of a randomly selected neutral image, followed immediately by another affective image, and so on. Images were sized 10.5 inches x 9.25 inches.

Data Manipulation and Cleaning

Reaction Intensity Quantification. The DART scoring process used in Study 2 was identical to the process used in Study 1. Two algorithms (*reaction start time* and *peak time*) were used to isolate emotional peak intensities.

Algorithms. The scoring algorithms were identical to those described in Study 1. These algorithms were applied to the raw data using a scoring program written in Python version 2.78.

Data Cleaning. We needed to verify and clean the algorithm output. To verify specific algorithm outputs (e.g., reaction start times) we randomly sampled and visually compared the Python outputs to time/emotion plots generated by Excel. The Python output closely matched the visual coding on 100% of the trials that were sampled.

The algorithms were not able to code both a reaction start time and a peak time on 4.68% of trials. Visual inspection of these trials suggested that the vast majority of uncodable trials were the result of subjects simply indicating no coherent reaction to the emotional images. Again, because emotion reactions are the focus of the present studies, we dropped these trials from the main analyses.

Results

Preliminary Models

Before addressing our key hypotheses, we created a single-purpose OLS data set that allowed us to examine relationship between individual differences in positive and negative emotional reactivity (peak intensities). These peak intensities were positively correlated across subjects, $r = .47$, $p < .001$. This correlation aligns with PF theory and suggests a common system (e.g., experiential openness) that underlies the generation of both positive and negative emotion.

Hypothesized Effects

Again, although we formulated our hypotheses with emotion reactivity influencing SWB, the present reactivity task forced us to model reactivity as the dependent variable and SWB as predictor. This modeling approach is of course casually arbitrary and thus does not influence our interpretation of the results.

Given the DART's nested data structure, MLM was the most sophisticated way of explaining variance in the dataset (Nezlek, 2008). To examine whether more intense emotion reactions were linked to higher SWB composite scores (level 2 individual difference), and also whether image valence (level 1 within-subject variable) moderated any such effects, we performed a MLM using SAS Proc Mixed (Singer, 1998). SWB scores were z-scored, and valence was coded as -1 for negative images and +1 for positive images ($M = 0$, $SD = 1$). To compare reactivity across valence, Peak Intensity scores associated with unpleasant stimuli were multiplied by -1.

Results of this model were quite similar to those of Study 1. The intercept of this model was significantly non-zero, $b = 291.16$, $t = 41.63$, $p < .001$, 95% CI = 277.36 to 305.24, indicating a moderate to strong average Peak Intensity. There was also a significant effect of valence, $b = -51.42$, $t = -11.57$, $p < .001$, 95% CI = -60.13 to -42.71, indicating that positive images triggered less intense reactions than negative images.

More importantly, this model revealed a significant level 2 main effect of the SWB composite on Peak Intensity scores such that greater Peak Intensity scores were associated with greater SWB levels, $b = 22.48$, $t = 3.20$, $p = .002$, 95% CI = 8.60 to 36.36. Estimated means revealed that happier people (+1 *SD* SWB) reported higher intensity reactions (*M* = 342.41), relative to unhappy people (-1 *SD* SWB). Also supporting our hypotheses, image valence did not significantly interact with this effect, *b* = 2.32, *t* = .53, *p* = .603.

Follow-up analyses replacing the SWB composite with FS or PE revealed the same hypothesized pattern of significant and non-significant effects. Both FS (b = 6.89, *t* = 3.14, *p* = .001, 95% CI = 9.36 to 36.62) and PE (b = 17.56, *t* = 2.53, *p* = .013, 95% CI = 3.81 to 31.31) positively predicted emotional intensity peaks and neither the FS effect (p = .958) nor the PE effect (p = .367) was associated with a significant valence interaction. Together, these results suggest that greater emotion reactivity may be linked to greater SWB.

DISCUSSION AND STUDY 3

Studies 1 and 2 provided evidence consistent with a PF-related perspective on SWB in the form of more pronounced situation-congruent emotion reactions. Study 3 is more process-oriented and seeks to model a core theoretical mechanism by which PF could lead to higher subjective quality of life. In particular, PF theory suggests that matching one's reactions to situations may allow the individual to behave more adaptively in life's continuously changing environments, and such adaptive behavior (e.g., coping) should then promote SWB (Kashdan & Rottenberg, 2010). Conversely, failing to respond to situational cues has been associated with decreased mental health (Peterson & Seligman, 1983) and it is not a coincidence that blunted emotion reactivity is a prominent correlate of depressed mood (Höflich et al., 2018). This idea that healthy emotion systems should show robust emotional and behavioral responses to the environment makes sense given that a central evolutionary purpose of emotional output may be to motivate organisms to adaptively alter behavior to match changing environments (Nesse & Ellsworth, 2009). A rabbit must flee when there is danger, and eat when there is food, and these reactions require drastically different motivations. In Study 3, we examine processes of this type using an implicit paradigm that asks people to make choices in a changing, emotional environment.

Method

Participants and General Procedures

With respect to a-priori power considerations, the Study 3 data structure and related statistical analyses examine effects that are very similar to those examined in Studies 1 and 2. Specifically, the key hypothesized effect in Study 3 is an MLM-based level 2 main effect. For this reason, sample size considerations in Study 3 were identical to those described in Study 1.

Participants (n = 98, M age = 19.41, SD = 4.1, 72.45% women, 85.71% Caucasian) were recruited and compensated via SONA. All laboratory facilities were as described in Study 1, and informed consent/IRB approval was obtained. Subjects completed an emotion and action task presented via E-Prime 2.0, and then completed a series of well-being assessments presented via MediaLab.

Subjective Well-Being Assessment

The SWB instruments used in Study 3 were the same ones used in Study 1. To obtain theoretically inclusive quantifications of both positive emotional experiences as well as broader judgments about the contents of one's life, we administered both the Flourishing Scale (Diener et al., 2010), or FS (M = 5.75, SD = .71, alpha = .88), and the positive component of the Scale of Positive Emotion and Negative Emotion (Diener et al., 2010) or PE (M = 3.70, SD = .60, alpha = .85).

As might be expected, we were interested in the relationship between adaptive behavior and overall subjective well-being, and a composite happiness score was created by standardizing and averaging the PE and FS scales (M = -.014, SD = .91). The FS and PE components were correlated at r = .63.

Hedonic Choice Assessment

Procedure Synopsis. This task was designed to model choices related to hedonic experience. When emotion systems are working properly, people should be motivated to approach pleasant stimuli and avoid unpleasant stimuli (Carver & White, 1994), which should affect their behavior. To model processes of this type in a somewhat implicit way, we created an environment that contained both a pleasant stimulus and an unpleasant stimulus and asked participants to choose where they would "want to be". On each of 50 trials, people were initially presented with two 4.2 x 3.1 inch IAPS images located at either the left or right of the computer screen. One image was always of pleasant valence and one image was always an unpleasant image. We first asked people to rate their state emotion while viewing each image. This was done to obtain some idea of how intense their emotional experiences were, but also to ensure that participants fully viewed and processed each image. Next, the primary task environment was presented. This environment (see Figure 2) presented these same two images as well as an avatar representing the participant. Participants were asked to choose a location between the two images where they wanted to "be". The instructions preceding the task were as follows:

> On this portion of the experiment (about 12 minutes), you will first see 2 slide images. You will be asked to rate the pleasantness of each slide. Next, you will see a spatial layout with "you" in the center and the images to your figure's left and right. Your task is to freely choose a location along the line that reflects where you would like to "be" for that trial. It is totally up to you where you decide to go along the line, including near the center or near one of the images.

Figure 2. Emotion and behavior task environment. Stimuli are included for descriptive purposes only- not actual IAPS images.

Affective Stimuli. Image selection procedures were similar to those used in Studies 1 and 2. Fifty images of each valence were first selected. These selections were made using IAPS SAM norm data for image pleasure and image arousal (Bradley & Lang, 1994). We first confirmed that the pleasant stimuli we selected (*M* pleasant = 7.38) were indeed more pleasant than the unpleasant images (*M* pleasant = 2.51), $F(1, 99) = 9560.43$, $p < .001$, PES = .99. To match positive and negative images for affective extremity, we created extremity scores for each image by calculating the absolute value of the pleasantness norm minus five (*M* pleasant = 2.38, *M* unpleasant = 2.49). Two ANOVAs were calculated to examine whether the pleasant and unpleasant image sets meaningfully differed by affective extremity, $F(1, 99) = 5.03$, $p = .027$, PES = .05, or arousal, $F(1, 99) = 0.21$, $p = .645$, PES = .00. There was a significant difference in valence extremity such that pleasant images were slightly less provocative than unpleasant images. However, the effect size was small (PES = .05), and main effects of valence where not of interest to our hypotheses. In fact, we expected a very large effect of image valence on avatar location choices, regardless of any difference in image extremity.

Trial Procedures. The task involved 50 trials containing 50 images of each valence. All trials had two major components: an image rating procedure, and an avatar placement procedure. On each trial, the two images first appeared on a black background. After a 2000 millisecond delay, subjects were asked to rate how "pleasant or unpleasant" each image was (1 = very unpleasant, 7 = very pleasant). Admittedly, this approach is not an ideal measure of reaction intensity, but it does reflect a state-like experience generated by a highly controlled (temporally and extremity) stimulus while the stimulus was still present. Trials always included a pleasant image and an unpleasant image, but the location of the images was randomly determined. Subjects rated the images before choosing where to place their avatar to ensure that the images were attended to and processed. Following image ratings, the emotional behavior environment was presented (see Figure 2). During this stage, the mouse cursor was replaced with the avatar icon so that the avatar could be freely controlled with the mouse. The trial ended when the subject moved the avatar to a location in the "black space" and clicked the mouse. There were 12 different possible locations for the avatar.

Emotional Flexibility Assessment. To quantify the extent to which people reported more extreme reactions to the provocative images, we subtracted the pleasantness rating made for the unpleasant slide from the pleasantness rating made for the pleasant slide (6 = pleasant slide is maximally pleasant and unpleasant slide is minimally pleasant, -6 = pleasant slide is minimally pleasant and unpleasant slide is maximally pleasant). The mean was 4.14 (*SD* = 2.13), indicating that people tended to have more pleasant reactions to positive relative to negative stimuli, but such scores also varied across trials and individuals. Similar valence difference scoring has been used successfully to link increased reaction intensity to well-being markers (e.g., Waugh et al., 2011).

Behavioral Flexibility Assessment. As we have said, well-calibrated organisms should adjust their behavior to approach rewards and avoid punishment. From this perspective, adaptive responding on the present task is reflected in the tendency to place one's avatar nearer to the slide rated as pleasant and further from the slide rated as unpleasant. To quantify the extent to which people placed their avatar in an adaptive location, each trial recorded the distance between subject's avatar placement and the pleasant image (-5.5 = furthest from the pleasant image, +5.5 = closest to the pleasant image). As expected, people tended to approach positive rather than negative stimuli (*M* = 3.79, *SD* = 1.70).

Results

Although we have formulated our hypotheses to align with existing theory, the present task forced us to model avatar choice as a dependent variable and SWB as predictor. This modeling approach is of course casually arbitrary and does not influence our interpretation of the results.

We used MLM in Study 3 because several of our variables are within-subject measures that were repeated many times. Studies 1 and 2 showed that people with more reactive or flexible emotion systems tended to be happier. Theoretically, one reason why such emotional flexibility could support greater well-being is that these emotional extremes motivate people to make situationally-congruent and adaptive choices. The aim of study 3 was to examine this possibility.

To replicate Studies 1 and 2, we first assessed whether Emotional Flexibility (here defined in terms of greater pleasantness in response to positive stimuli and greater unpleasantness in response to negative stimuli) is linked to SWB in this new paradigm. To evaluate this question, we ran an MLM via SAS Proc Mixed (Singer, 1998) that assessed links between Emotional Flexibility (dependent variable) and SWB composite scores (level 2 independent variable). SWB scores were z-scored (Aiken & West, 1991). Results of this model were quite similar to those of Studies 1 and 2. As expected, this model revealed a significant level 2 main effect of the SWB composite on Emotional Flexibility scores such that greater Emotional Flexibility was associated with greater SWB levels, $b = .39$, $t = 3.83$, $p < .001$, 95% CI = .19 to .59.

Next, we were interested in whether this greater emotional flexibility would also be linked to situation-congruent responses in the form of greater Behavioral Flexibility. Accordingly, we ran an MLM with Emotional Flexibility (level 1, repeated variable) predicting Behavioral Flexibility (level 1 repeated variable). Emotional Flexibility scores were person-centered to control for individual differences in average reactivity (Enders & Tofighi, 2007). Here, Emotional Flexibility Scores were, as expected, linked to greater Behavioral Flexibility, $b = .39$, $t = 13.51$, $p < .001$, 95% CI = .34 to .45. What this means is that the average person displayed larger approach-avoidance effects when they experienced larger differences in feelings toward the two stimuli.

Next, we evaluated the novel (in the present paper) question of whether behavioral flexibility would itself be linked to SWB levels. This was indeed the case—a third MLM revealed that greater Behavioral Flexibility (dependent variable) was robustly linked to greater SWB (level 2 independent variable), $b = .33$, $t = 3.75$, $p < .001$, 95% CI = .15 to .50. This model suggests that, in this dataset, 1 *SD* of change in well-being predicted average placement .33 units closer to the more pleasant image (which represents a 5.5% closer to the positive image or 5.5% further from the negative image). Here, we should remember that there were *no* constraints whatsoever on avatar placement and behaviors following feelings would reflect intuitive or spontaneous behavior.

DISCUSSION AND STUDY 4

If we allow it to be true that nearness to pleasant stimuli and distance from unpleasant stimuli is beneficial to well-being, then one of the simplest interpretations of the present findings is that people who are more emotionally reactive to stimuli seem abler to accomplish this kind of basic approach-avoidance task. That is, they seem abler to avoid unpleasant things and approach pleasant things, and this ability, if externally generalizable, should improve one's quality of life. These findings are regressive/correlational, and future research should address such processes using manipulation-based approaches.

Study 4 returns to our original question: the relationship between emotion reactivity and SWB. Because the previous studies were all laboratory based, it is reasonable to wonder whether the DART has predictive value for daily well-being. Thus, Study 4 employed a similar DART paradigm and examined whether greater Peak Intensity on this task would predict daily measures of SWB up to four weeks later.

Method

Participants and General Procedures

Study 4 was a two-part study including an initial laboratory portion followed by a 14-day daily diary protocol. General laboratory procedures were very similar to those described in Study 2. The lab portion was pre-scheduled to run for two consecutive weeks and recruited a college-aged convenience sample (n = 128, M age = 19.08 SD = 2.99, 63.28% women, 88.28% Caucasian). The laboratory procedures for Study 4 included only demographics and the continuous emotion rating task. The daily protocol began assessing daily well-being levels one week after the lab sessions were complete and continued for 14 consecutive days.

To estimate the sample size required for acceptable power in this study, we followed procedures designed for a-priori power analysis of multi-level models (described above). The primary question in Study 4 was whether Peak Intensities would predict daily well-being levels. Trial-level individual differences in intensity were averaged across trials, for each person, to create single level 2 predictors for each participant.

Dynamic Affect Reactivity Task

Procedure Synopsis. In a randomized order, participants viewed 10 pleasant and 10 unpleasant images. These images were selected from the International Affective Picture System or IAPS (Lang et al., 1997). After viewing a "get ready" image for three seconds, provocative images were presented for 5 seconds (see Figure 1 above). After this, the slide was replaced with a blue screen containing instructions to continue monitoring and rating one's feelings. Participants rated their emotions using a computer mouse. Pre-task instructions were as follows:

> During this section (about 7 minutes), we would like you to report your EMOTIONAL REACTIONS TO DIFFERENT IMAGES. We are interested in your basic feelings, that is how pleasant you feel or unpleasant you feel, in response to the image.
>
> To rate your feelings, you will use a rating scale (that will appear to the right on your screen) and the COMPUTER MOUSE.
>
> Try your best to rate your reaction to the image EVEN AFTER IT HAS DISAPPEARED. After the image disappears, continue to monitor and rate your emotional reaction to the slide FOR THE FULL DURATION OF THE TRIAL!

Monitoring and rating your feelings is a strange experience at first. However, human emotions are ALWAYS CHANGING, so try to notice and report these changes.

Affective Stimuli. Image selection procedures were similar to those used in Study 1. Ten images of each valence/pleasantness were first selected. Selections were made using IAPS SAM-based norm data for image pleasure (1 = "very unhappy" experience while viewing, 9 = "very happy" experience while viewing) and image arousal (1 = a sleepy looking manikin/avatar, 9 = a manikin with an exploding stomach; Bradley & Lang, 1994). We first confirmed that the pleasant stimuli we selected (*M* pleasant = 7.07) were indeed more pleasant than the unpleasant images (*M* = 2.96), $F(1, 18) = 2570.21$, $p < .001$, PES = .99. To match positive and negative images for affective extremity, we created extremity scores for each image by calculating the absolute value of the pleasantness norm minus five (*M* pleasant = 2.07, *M* unpleasant = 2.04). Two ANOVAs were calculated to ensure that the pleasant and unpleasant image sets did not significantly differ by affective extremity, $F(1, 18) = .14$, $p = .716$, or arousal, $F(1, 18) = .69$, $p = .416$. Pleasurable affective images featured enjoyable scenes including cute animals or boating, and unpleasant images featured disturbing scenes such as physical attacks, car accidents, or crime scene photos.

Trial Procedures. Participants were randomly assigned to either the "top good" condition or the "top bad" condition. The task included 20 trials. Images were sized 10.5 inches x 9.25 inches and presented for 5 seconds followed by a 15 second blank screen. Rating bar locations were recorded every 100 milliseconds and ranged from -500 to +500, depending on the participants' positioning of the rating bar.

Data Cleaning and Variable Quantification

Algorithms. The scoring algorithms were identical to those described in Study 1. These algorithms were applied to the raw data using a scoring program written in Python version 2.78 for this purpose.

Data Cleaning. We again need to verify and clean the algorithm output. To verify specific algorithm outputs (e.g., reaction start times), we randomly sampled and visually compared the Python outputs to time/emotion plots generated by Excel. The Python output closely matched the visual coding on 100% of the trials that were sampled.

The algorithm was not able to code both a reaction start time and a peak time on 8.7% of trials. Visual inspection of these trials suggested that the vast majority of uncodable trials were the result of subjects simply indicating no coherent reaction to affective images. We dropped these trials from the main analysis.

Emotion Intensity Quantifications. The scoring process for Peak Intensity scores was identical to the processes described in Studies 1 and 2. The two previously described algorithms were used to determine the Peak Intensity of each trial's emotion reaction. To allow for the direct comparison of Peak Intensity effects between valences, we multiplied all Peak Intensity scores associated with negative trials by -1. Next, to use these Peak Intensity scores as predictors of daily SWB, we needed to collapse across trials to calculate two average scores for each person. First, an *Average Peak Intensity* score was created by averaging all Peak Intensity scores across both pleasant and unpleasant trials (M = 260, SD = 76). Second, to quantify the effect of image valence on Peak Intensity scores, we subtracted each person's average Peak Intensity scores on negative trials (M = 317, SD = 100) from average Peak Intensity on pleasant trials (M = 203, SD = 91) to create a single difference score (M = -114, SD = 116), with relatively higher numbers indicating greater relative positivity. This score is subsequently referred to as referred to as *Average Intensity Difference*.

Daily Protocol

After the laboratory portion was complete, the daily diary phase began. Participants received email reminders with Qualtrics survey links at 7 p.m. on each of 14 consecutive days. To limit retrospection, these links remained active for exactly 12 hours. To increase the power and validity of the results, subjects completing less than 9 surveys were excluded from the study (West, Ryu, Kwok, & Cham, 2011). The final dataset included 1607 daily reports.

Daily Satisfaction. Thus far, we have relied on the Flourishing Scale (Diener et al., 2010) to assess global quality of life judgments. However, the FS assesses multiple life domains that are quite specific, and each of these domains may not be relevant on a given day. Thus, this scale did not lend itself well to repeated daily assessment. Instead, we replaced the FS with more general measures of daily life satisfaction, which is also a commonly used subjective judgment of one's quality of life. The daily satisfaction measure is much broader than the FS and so is more suited to daily fluctuation. We did, however, desire some heterogeneity in the assessment domains, and we therefore obtained satisfaction reports of 2 key life domains: the self, and one's overall "health". Each day, participants rated the extent to which they felt (1 = not at all, 5 = extremely) "satisfied with myself" (M = 3.30, SD = 1.04) and "satisfied with my health" (M = 3.28, SD 1.09).

Daily Positive Emotionality. The Positive Emotionality component of the SPANE (Diener et al., 2010) was altered to target daily emotions. Each day, participants indicated the extent to which (1 = not at all; 5 = extremely) they felt each of 3 wide-ranging, non-specific positive feeling states (Positive, Happy, Excited). Item level scores were averaged for each day (M = 3.60, SD = .95, alpha = .92).

Daily Negative Emotionality. Although we did not have hypotheses about the presence or absence of ill-being markers, we included Negative Emotionality as a contrast variable. For this purpose, the Negative Emotionality component of the SPANE was altered to target daily emotions. Each day, participants indicated the extent to which (1 = not at all; 5 = extremely) they felt each of 3 wide-ranging, non-specific negative feeling states (Sad, Negative, Distressed). Item level scores were averaged for each day (M = 1.88, SD = .86, alpha = .86).

Results

Study 4 assessed daily outcome variables that were measured repeatedly within-subjects. Such data should be analyzed using MLM (Nezlek, 2008). In this results section, we modeled links between the 2 DART predictors (Average Peak Intensity and Average Intensity Difference) and 4 daily SWB outcomes (Satisfaction Self, Satisfaction Health, Positive Emotionality, and Negative Emotionality). We ran a series of “level 2 main effect” MLMs using SAS Proc Mixed (Singer, 1998). In all of these models, the DART predictors were z-scored, and our outcomes retained their original units (Aiken & West, 1991).

Average Peak Intensity Predicting Daily SWB

As hypothesized, these models revealed a positive relationship between Average Peak Intensity and each of the well-being indicators: Satisfaction with Self, $b = .147$, $t = 2.26$, $p =$ .026, 95% CI = .02 to .27; Satisfaction with Health, $b = .159$, $t = 2.21$, $p = .029$, 95% CI = .02 to .30; and Positive Emotionality, $b = .166$, $t = 2.78$, $p = .006$, 95% CI = .05 to .28. However, DART reactivity was not predictive of Negative Emotionality, $b = -.014$, $t = -.26$, $p = .799$, 95% CI = -.12 to .09. Overall, the models suggest that greater psychological health is associated with having intense reactions to both positive and negative stimuli. This finding is consistent with our expectations as well as with theories describing the positive relationship between psychological flexibility (PF) and happiness or well-being.

Average Intensity Difference Predicting Daily SWB

We next turned to the valence-specific score (i.e., having more intense reactions to positive than negative stimuli, relative to other people) to see whether it could also predict daily WB. This Average Intensity Difference score was not predictive of positive emotionality, $b = .083$, $t = 1.36$, $p = .177$, 95% CI = -0.04 to .20. Somewhat surprisingly, the two Daily Satisfaction models showed that larger Average Intensity Differences were significantly predictive of both Daily Satisfaction with Self, $b = .200$, $t = 3.11$, $p = .002$, 95% CI = .07 to .32, and Daily Satisfaction with Health, $b = .255$, $t = 3.67$, $p < .001$, 95% CI = .12 to .39, suggesting that more positive (and/or less negative) average peak intensities were predictive of greater daily SWB. These results were surprising because reaction intensity effects on SWB in the lab were not moderated by valence.

Average Intensity Differences were also marginally predictive of less negative emotionality, $b = -.090$, $t = -1.7$, $p = .091$, 95% CI = -.19 to .01, such that more positive overall reaction intensities (greater difference score) could be a predictor of less negative emotionality in daily life.

To better understand these significant and marginal effects of valence, we ran a series of follow-up MLMs with Average Peak Intensity scores separated by valence (and standardized). These models revealed that more intense reactions to *positive* DART stimuli were not predictive of daily Negative Emotionality, $b = -.068$, $t = -1.3$, $p = .196$, 95% CI = -.17 to .04, but were predictive of greater Satisfaction with Self, $b = .249$, $t = 4.01$, $p < .001$, 95% CI = .13 to .37, and greater Satisfaction with Health, $b = .295$, $t = 4.35$, $p < .001$, 95% CI = .16 to .43.

Average Peak Intensity reactions to negative stimuli were not predicative of any daily well-being markers: Negative Emotionality, $b = .041$, $t = .78$, $p = .430$, 95% CI = -.06 to .15, Satisfaction with Self, $b = -.003$, $t = -.04$, $p = .966$, 95% CI = -.13 to .12, or Satisfaction with Health $b = -.026$, $t = -.35$, $p = .739$, 95% CI = -.17 to .12. Taken together, these results suggest that although both positive and negative reaction intensity are significantly predictive of daily well-being (as seen in the main effects models above), the effect of positive intensity may be stronger than the effect of negative intensity.

GENERAL DISCUSSION

In this project, we hypothesized that more intense emotion reactions would predict greater subjective well-being. We found support for this hypothesis in three studies. More intense responses on the DART predicted greater SWB reported in the lab as well as greater well-being in daily life. It was also hypothesized that one potential mechanism for this intensity/SWB effect would be a link between more intense emotion reactions and more flexible, situation-congruent behavior. We find support for this hypothesis in Study 3, where people who reported more intense emotion reactions were more likely to avoid objectively unpleasant stimuli, and to approach pleasant stimuli. Importantly, both of these behavioral tendencies were also predictive of higher SWB levels. These results provide evidence supporting the functionality of flexible emotion reactivity. In the General Discussion, I focus on implications of the present findings and I discuss important future directions for the research.

Methodological Implications

The present results are significant because they suggest that the DART, if further validated, may represent a methodological advancement in affective assessment. The DART is the first paradigm capable of both 1) measuring subjective emotion reactions over time with high temporal resolution and 2) coding millions of these affective data points to determine the precise intensity of individual emotion reactions that are occurring at different times on different trials. When the DART is further developed, this high resolution coupled with the coding algorithms should enable researchers to operationalize numerous additional elements of emotion reactions (e.g., reaction latency, rise time, peak duration, reaction duration, drop time, etc.). These capabilities are scientifically useful because they allow one to measure emotion reaction intensities without relying on questionnaire reports that may confound reaction intensity with unrelated constructs such as reaction duration (Davidson, 1998). The DART can assess a given person's typical reaction intensity while taking into account intra- and inter-individual differences in the precise time-point at which a given emotion reaction reaches its peak intensity. This flexibility allows the measurement of reaction intensity without arbitrarily choosing a specific time-point at which to assess state emotion intensity. Such a process is useful because different people's emotion reactions reach their peak intensity at different times (Davidson, 1998, 2000). For instance, a high STAI emotion rating could be a correct reflection of an intense emotion reaction, or, if the STAI is inadvertently administered after most people have reached their peak emotion, a high STAI rating could reflect a tendency to have elongated emotion reactions (Davidson, 1998).

In terms of encoding continuous streams of emotion data, the DART algorithms can also determine a given stream's peak intensity without relying on an average calculated during the normative peak reaction window/epoch. Such a procedure has often been used in emotion reactivity research (e.g., Mauss et al., 2005), and this averaging process cannot control for individual differences in peak latency (Luck, 2014). The DART algorithms can also assess peak intensity within a continuous stream of emotion data without using a simple "max" value that can be noisy, error prone, and unreliable (Luck, 2014). In sum, the DART procedures could mitigate several problems with existing measures of emotion reaction intensity.

The present results are also important because they suggest that the methodological advantages of the DART may provide some clarity on key questions in affective science. Specifically, the DART can help address the existing uncertainty about the relationship between emotion reactivity and well-being. It is worth taking a moment to add some detail to my previous discussion of methodological challenges faced by the emotion dynamics literature and how the

DART can address these challenges. In the emotion dynamics literature (which studies emotion reactivity patterns over time), there is a reasonable consensus that elongated emotion reactions are predictive of lowered well-being. However, there is significant disagreement about the impact of emotion *intensity* on well-being. This literature has two notable “camps”. As I have said, each of these camps tends to use its own general set of emotion intensity assessment methods, and these camps come to opposing conclusions about the relationship between reaction intensity and well-being. In the first camp, there is research modeling discrete emotion reactions to controlled stimuli using methods that obtain high temporal resolution. One example of such methods is research using a combination of stimulus onset asynchrony (SOA) coupled with discrete probes of emotional responding (e.g., Waugh et al., 2011). This research has generally shown a *positive* relationship between indicators of reaction intensity and well-being (Schaefer et al., 2018; Waugh et al., 2011).

Second, there is a group of emotion dynamics researchers using state (e.g., EMA) or summary assessments that capture emotion reactivity to stimuli in daily life. This literature using daily EMA-related techniques (which is the more common approach) has reported *negative* associations between reaction intensity and well-being (Houben et al., 2015). I believe that these two literatures diverge in part because of the methodological limitations that I have alluded to. In particular, these daily/EMA studies generally do not collect information about activating stimuli (whether with respect to intensity or frequency), or the stimulus timing relative to the emotion intensity measurements. To help explain the advantages of the DART, I will briefly explain several specific flaws in this Daily/EMA approach and how the DART solves these problems.

Daily emotion/well-being studies have generally operationalized individual differences in the tendency to experience intense emotion reactions as a standard deviation of multiple EMA state emotion measures collected over time (Houben et al., 2015). Within-person SD over time is clearly a reliable measure of some construct (given that it reliably predicts well-being in multiple studies). However, the most obvious problem with SD as a measure of emotion intensity is that people with lowered well-being tend to experience more aversive life experiences, which engender more negative emotion reactions (Argyle, 2013; Charles et al., 2013). For such participants, SD calculated across a random sample of state emotions could therefore be larger not because these people tend to have more intense emotion reactions but because they tend to have more frequent emotion reactions. Given this confound, some EMA studies have controlled for activating event frequency.

However, a second problem with this daily approach is equally significant but subtler. As I have said, this EMA literature has often operationalized individual differences in reaction intensity in terms of larger standard deviations of state affect collected over time (see the Houben et al., 2015 meta-analysis). I will argue that this measurement approach is significantly flawed. In particular, a person with a tendency to experience intense emotion reactions might show a *smaller* standard deviation if those reactions tend to be of shorter duration. Such a person could show a smaller SD simply because they spend more time at emotional baseline. In contrast, a person with *less* intense but longer emotion reactions may show a *larger* standard deviation because they spend less time at baseline. I illustrate this effect in Figure 3, which shows the affect intensity over time of a hypothetical person who has intense but short-lived reactions (top panel) and a person with blunted but longer reactions (bottom panel).

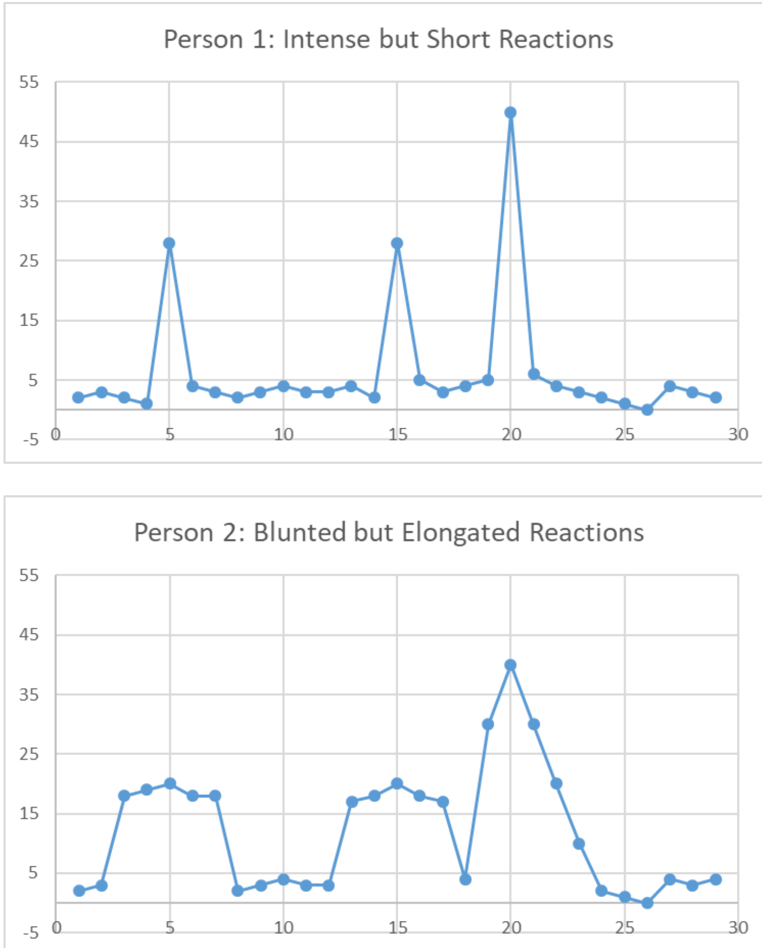

Figure 3. Hypothetical state affect intensities of 2 prototypical reactivity profiles. The top panel depicts a person showing emotion reactions that are intense but short, and the bottom panel depicts a person showing a blunted but elongated pattern.

If one were to randomly sample the state affect of these two people, the person with blunted but elongated reactions might show a larger standard deviation than the person with more intense reactions. This is obviously a problem when larger standard deviations are often operationalized as reflecting more intense reactions. To verify that a person with intense but short-lived reactions could actually show a smaller SD, I used a random number generator to select 7 affect snapshots from each of the displayed individuals (I did **not** test multiple samples of the distribution until I found a good result. I took one sample). As I expected, Person 1 had a smaller SD of these 7 samples (9.31), relative to Person 2 (10.86). Given that Person 1 has more intense emotion reactions (in reality; see Figure 3), this demonstration suggests that SD is a flawed emotion intensity metric. This problem with SD is also true of the related mean-subsequent-squared-difference calculation (MSSD) which has also been used in emotion dynamics/well-being research. This discussion is crucial because the emotion dynamics literature has used SD/MSSD frequently in its examination of intensity/well-being relationship (for a meta-analysis, see Houben et al. 2015). Importantly, research that accounts for stimulus timing, frequency, and intensity (e.g., the DART, startle-based research) has often show opposite intensity/well-being relationships as this SD research.

In summary, commonly employed measures of emotion reaction intensity may be somewhat flawed. This is true of both conventional self-reports of emotion reactivity (e.g., state affect following a provocative stimulus), and standard deviation approaches used in daily or EMA research. The DART procedures can address some of these flaws by administering controlled stimuli and measuring emotion intensity with high temporal resolution. Given these flaws and the relative sophistication of the DART procedures, the present results call into question the conventional idea that greater reaction intensity is linked to lower well-being.

Theoretical and Applied Implications

The present findings (and those of other related studies) may cause affective researchers to rethink our current understanding of the relationship between emotion reactivity and well-being. Such an alteration could impact our theoretical understanding of the causes of mental well-being, as well as therapeutic approaches to helping people who struggle with affective disorders. Naturally, the scientific consensus on whether intense emotion reactions are healthy or unhealthy should influence approaches to treating affective disorders (Davidson, 1998, 2000; Kuppens & Verduyn, 2017). Even outside of clinical contexts, lay views of the healthy or unhealthy nature of intense emotion reactions also shape popular approaches to emotion regulation and the ubiquitous search for happiness. For example, viewing intense emotion reactions as unhealthy may contribute to the prevalence of regulation approaches such as suppression or denial that have been shown to be significantly less effective (Gross, 2015). Below, I discuss some ways in which our results matter in these contexts.

The present results suggest that flexible, situation-congruent emotion reactivity is healthy and likely supports adaptive choices. These results align with the psychological flexibility theory of mental well-being (Kashdan & Rottenberg, 2010) and begin to extend this theory into the domain of emotion reactivity. PF theory argues that a cornerstone of mental health is the ability to adjust psychologically and behaviorally to changing environments, and our results suggest that situation congruent emotion reactivity may be one key feature of such a flexible system. If emotional flexibility is a feature of psychological flexibility, the present results also suggest that interventions that have been shown to increase psychological flexibility could be viable approaches to increasing emotional flexibility. Examples of such interventions include acceptance-oriented trainings (Hayes et al., 1999; Hayes et al., 2012; Miller, Rathus, & Linehan, 2006) and mindfulness oriented trainings (Kangasniemi, Lappalainen, Kankaanpä, & Tammelin, 2014; Masuda & Tully, 2012).

In addition, the present results have mechanistic implications. The results point to adaptive behavior/decision making as one mechanism by which robust emotion reactivity may increase quality of life. In a basic sense, emotion reactivity can be thought of as a homeostatic process (Craig, 2003), and homeostasis fundamentally underlies the activity/behavior of all organisms (Craig, 2003; Damasio & Carvalho, 2013). Organisms must possess a mechanism that motivates the creature to modify its cellular/bodily activity to match changing environments—

that is, to be flexible. To survive, for instance, creatures must consume in the presence of food and flee in the presence of danger (Nesse & Ellsworth, 2009). Study 3 speaks to process-oriented views of the link between reaction intensity and well-being. The findings support theories suggesting that flexible emotion reactivity benefits people (and supports well-being) largely because reactivity motivates people to behave adaptively in changing environments (Craig, 2003; Damasio & Carvalho, 2013; Kashdan & Rottenberg, 2010; Waugh et al., 2011).

If the present results are valid, and it is true that more intense emotion reactions are not necessarily problematic for our mental health, this finding suggests that *other* emotion reactivity factors are key to understanding the apparent relationship between emotion reactivity and lowered well-being. For instance, given the mixed results for the intensity/well-being link as well as the consistent results showing a negative relationship between reaction duration and well-being (Houben et al., 2015), it may be that a tendency to experience elongated reactions is a key feature of maladaptive emotion systems **(**Derryberry & Reed, 2002; Schaefer et al., 2013; Taubitz et al., 2013). Such a conclusion would have implications for emotional disorder treatment. For instance, viewing elongated emotion reactions as a key contributor to lowered well-being suggests (as does psychological flexibility theory itself) the importance of mind states that have been specifically associated with shortened emotion reactions. Examples of such mind states include acceptance/openness/nonattached attitudes (Klein & Robinson, in progress; Low, Stanton, & Bower, 2008). Such acceptance-oriented mind states have been the focus of a wave of new psychotherapies (e.g., Dialectical Behavior Therapy; Miller et al., 2006). The present results, as well as the emotion dynamics research writ large, seem to suggest that a therapeutic emphasis on developing acceptance-related mind states could be a powerful approach to increasing well-being levels.

Although there appear to be important flaws in some emotion reactivity measures, and that based on these flaws we should revisit the conclusions of research suggesting that intense emotion reactions are generally unhealthy, it also seems likely that typical self-reported emotion reactivity measures are capturing *some* valid variance in reaction intensity. Thus, there may indeed be some link between more intense emotion reactions and lower well-being. One possible explanation for the discrepancy between the present research and research showing the opposite reactivity/well-being patterns (e.g., Gross et al., 1998) could be that many of the relevant studies are specifically examining emotion reactions to negative stimuli (see Houben et al., 2015). Given such ideas, one could interpret the overall pattern of results in the emotion reactivity literature as an argument that more intense emotion reactions to *negative* stimuli have been associated with both higher well-being (e.g., the present main effects of reaction intensity) and lower well-being (e.g., the negative emotion specific findings: Gross et al., 1998). Thus, a key difference between individuals with high SWB and those with high ill-being markers like neuroticism is the tendency to report intense reactions to positive stimuli. Such a conclusion aligns with prominent research teams who have argued for the well-being benefits of robust emotion reactivity to positive stimuli (Ong, Bergeman, Bisconti, & Wallace, 2006; Tugade, & Fredrickson, 2004). In this work, the well-being benefits of positive emotion reactivity seem to be substantial.

Given these ideas, the mechanisms underlying positive emotion reactivity become quite interesting, and it would seem important to investigate how we might increase positive emotion reactivity. Here, attitudes related to openness and acceptance again become salient avenues for intervention. If we think of positive emotion as being principally caused by goal/need congruent appraisals (which is our current understanding of the matter; Moors, Ellsworth, Scherer, & Frijda, 2013), one potentially fruitful intervention could be to teach people to adopt more expansive and permissive goals/needs—that is, to teach people to adopt goals that are more realistic and easily met. One could adopt, for instance, a specific goal/or need for one's life to live *more or less* stress free rather than completely stress-free. Such inclusive or flexible attitudes directed toward the inevitable vicissitudes of life have consistently predicted enhanced happiness and well-being (Sahdra, Shaver, & Brown, 2010; Sahdra, Ciarrochi, & Parker, 2016).

In sum, the idea that lowered well-being is driven simply by an overactive negative emotion generation system is possibly too simplistic. One could argue that a mark of psychopathology and lower well-being, in terms of reaction intensity, may be an *imbalance* between the BIS and BAS systems characterized by an active negative emotion generation system paired with an inactive positive emotion generation system.

A Conception of Duration as Reactivity

As I have said, there are empirical reasons (including the present data) to believe that the emotion reactions of mentally healthy people are more intense (e.g., Waugh et al., 2011) and also shorter (e.g., Houben et al., 2015). Currently, there is no direct theory to explain such a phenomenon. I propose that a tendency to generate intense but brief emotion reactions could be a manifestation of a single, flexible, well-calibrated emotion *reactivity* system. Here, we can think of "emotion reactivity" as an adjustment (increase or decrease) in emotion output in response to the changing environment. As I have discussed, change is a defining feature of all environments—all events that begin will inevitably end. Therefore, it can be considered a kind of natural law that unpleasant/aversive stimuli will eventually cease and these negative stimuli must be replaced by either a neutral or a positive stimulus. Given this, an emotion system that tends to produce shorter emotion reactions following the offset of a provocative event could simply be a system that is generating a robust "reaction" to the neutral environments that often follow the event. I would argue that more intense emotion reactions on the DART are an indicator of this kind of flexible, well-calibrated emotion generation system, and that this system supports psychological well-being.

Such a perspective helps us understand both the present findings as well as the body of research consistently linking shorter emotion reactions to higher well-being. In this way, longer duration emotions (and their association with decreased well-being) may be partially rooted in emotion reactivity systems themselves—emotion systems that fail to detect/process or react to the neutral (or positive) environments that eventually follow negative stimuli. To my knowledge, this is a novel perspective—I am suggesting that shorter emotion reactions may be a key symptom of a healthy emotion *reactivity* system.

Moreover, if we reverse perspectives and think about elongated emotion reactions as being partially caused by blunted reactivity systems, this duration-as-reactivity perspective would also help explain the counterintuitive pattern of blunted yet elongated stress/emotion reactivity patterns typical of the chronically stressed (Sapolsky, 2004) and the depressed (Taubitz et al., 2013). The emotion reactivity systems of the chronically stressed, for instance, do not seem to produce intense stress reactions in response to aversive events, and these systems also do not seem to flexibly deactivate when stressors are no longer present.

This duration-as-reactivity perspective is also intriguing because it could provide a further explanation for links between well-being and situational awareness constructs like mindfulness. If one remains in psychological contact with the present moment, which is the essence of a mindful state (Brown & Ryan, 2003), one's emotion generation system should be more likely to attend to and flexibly react to positive, negative, and neutral stimuli. Such flexible reactivity should produce intense but short emotion reactions, and all the previously discussed psychological benefits associated with such reactivity patterns. In this way, links between mindfulness and emotional well-being may be partially mediated by situation-congruent emotion reactivity.

Future Research

Future research should continue to validate the DART as a measure of emotion reaction intensity. Such efforts would be challenging for at least two reasons. The DART assesses streams of subjective emotion over time and codes this output into constructs that represent key dimensions of emotion reactivity. The most obvious problem with the DART is that existing measures of subjective reaction intensities do not achieve the temporal resolution required to specifically isolate reaction intensities from, for instance, reaction durations (Davidson, 1998). As I have said, existing summary retrospective measures like the Self-Assessment Manikin or state measures like the STAI probably capture multiple reactivity elements such as reaction intensity, reaction duration, and reactivity threshold. Thus, to validate the DART, we may require emotion intensity measures with higher temporal resolution.

A second difficulty in further validating the DART is that the subjective reactivity measured continuously over time may not correlate with other emotion response systems that can be measured at high temporal resolution. For example, Mauss and colleagues (2005) used a continuous indicator of subjective emotion state and showed that this indicator did not correlate with continuously assessed physiological emotion markers like heart rate variability. This lack of correlation between subjective and physiological emotion components seemed true even when controlling for differences in response timing/lag between the different components (Mauss et al., 2005).

One possible approach to solving these problems could be to correlate DART intensity with affect-modulated startle responses. Such a startle-magnitude approach to measuring emotion reaction intensity captures more behavioral markers of emotion reactivity, and such markers have been shown to be more closely related to subjective intensity over time (Mauss et al., 2005). Affect-modulated startle response measures could also be viable given that they can be paired with an SOA technique to achieve strong temporal resolution of emotion intensity over time (Schaefer et al., 2013; Waugh et al., 2011).

In addition to validation, future research should continue to expand the capabilities of the DART. Ideally, the DART would be capable of assessing individual differences in multiple dynamic components of emotion reactivity. In terms of research examining links between emotion reactivity and well-being, the DART could be used to assess the relationship between reaction duration and psychological well-being. The present DART paradigm had the potential to assess reaction duration, but when the algorithms designed to assess duration where applied, these algorithms yielded too many missing data points. In particular, the failure to capture reaction duration was principally caused by excessive missing data (>25%) when applying algorithms to code the time points at which peak emotion intensities began returning to baseline. Visual inspection of the emotion streams associated with these missing data suggested that the vast majority of the missing data points were the result of emotion reactions that peaked, but did not begin to return to baseline during the 10-15 second window afforded. In my opinion, future DART paradigms should incorporate trials that are at least 25 seconds that also include neutral images presented during the recovery period. Such procedures should result in a higher percentage of codable reaction durations.

Future research should also focus on improving the DART procedures to assess individual differences in reaction threshold. Reaction threshold is a core aspect of dynamic perspectives on emotion reactivity (Davidson, 1998; Davidson, 2000; Evans & Rothbart, 2007). Reaction threshold refers to the intensity of stimulation required to illicit a significant change in emotion output (Davidson, 2000). Lower reaction thresholds are associated with increased frequency of negative emotion reactions (e.g., Evans & Rothbart, 2007) and lowered well-being (Scherer, 2000). Thus, individual differences in reaction thresholds should be an important contributor to mental health and well-being (Davidson, 1998), and the DART should be uniquely capable of assessing these differences. In a repeated-measure paradigm like the DART, individual threshold values could be determined using Signal Detection Theory methods (SDT: Green & Swets, 1966). Reaction thresholds would be analogous to SDT's "bias" construct. In SDT theory, a large bias value represents an individual's tendency to report a signal when there is none. Using preexisting image norms, the DART stimuli could be selected to include neutral, very low extremity, and medium extremity images of each valence. SDT bias formulas could then be applied to DART reaction intensity (or duration) values to determine individual differences in a tendency to show meaningful emotion reactions to neutral or low-intensity images. We could then use these biases to predict various markers of happiness and well-being.

Finally, the present findings are discussed in terms of emotion intensity predicting or impacting SWB. However, our studies are mostly correlational, and there is limited extant evidence for reaction intensity preceding well-being. One reason for the present interpretation is that viewing the relationship from the opposite direction (well-being predicts intensity) leads to predictions that are not aligned with our results. For instance, state negative affect tends to potentiate reactivity to threatening stimuli (Kaviani et al., 2004), and state positive affect is associated with enhanced reactivity to desirable stimuli (Catalino & Fredrickson, 2011). Thus, one reason that happy people would report more intense reactions to positive DART images might be because they are experiencing greater positive affect at the time. However, this potentiation perspective would predict that unhappy people should show greater reactivity to negative stimuli, but we show no such interaction. In fact, the lack of a significant interaction in this (and other) studies of emotion intensity/well-being is perhaps the most useful result of the present work.

Naturally, future research should attempt to examine a potentially causal relationship between emotion intensity and well-being. One challenge for this research would be that individual differences in emotion intensity may be difficult to manipulate given that they could be partially reflect a temperamental feature of the person (Evans & Rothbart, 2007). Given this, one plausible means of establishing the temporal precedence of emotion reactivity would be to use a T1/T2/T3 longitudinal study. Such a study could capture within-person fluctuations in DART emotion intensity that naturally occur over longer time periods and link changes in these measures to future within-person changes in SWB.

CONCLUSIONS

If further validated, the DART may represent a meaningful advancement in affective assessment. This advancement would originate from the DART's ability to record and code high-resolution streams of subjective affective changes induced by controlled stimuli. The DART capabilities could help to clarify important questions facing affective science such as the relationship between emotion reaction intensity and well-being. This question is not simply one of basic science—the true relationship between reaction intensity and mental well-being has key implications for interventions targeting psychological health.

REFERENCES

Aiken, L. S., and West, S. G. (1991) *Multiple regression: Testing and interpreting interactions*, Thousand Oaks, CA: Sage.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC:

Argyle, M. (2013). *The psychology of happiness*. New York, NY: Routledge.

Barlow, D. H., & Kennedy, K. A. (2016). New approaches to diagnosis and treatment in anxiety and related emotional disorders: A focus on temperament. *Canadian Psychology/Psychologie Canadienne*, *57*, 8–20.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*. New York, NY: Guilford Press.

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *84*, 822.

Browne, W. J., Lahi, M. G., & Parker, R. M. (2009). *A guide to sample size calculations for random effect models via simulation and the MLPowSim software package*. Bristol, United Kingdom: University of Bristol.

Bryant, F. B. (2003). Savoring beliefs inventory (SBI): A scale for measuring beliefs about savouring. *Journal of Mental Health*, *12*, 175–196.

Carl, J. R., Fairholme, C. P., Gallagher, M. W., Thompson-Hollands, J., & Barlow, D. H. (2014). The effects of anxiety and depressive symptoms on daily positive emotion regulation. *Journal of Psychopathology and Behavioral Assessment*, *36*, 224-236.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319-333.

Catalino, L. I., & Fredrickson, B. L. (2011). A Tuesday in the life of a flourisher: The role of positive emotional reactivity in optimal mental health. *Emotion*, *11*, 938.

Charles, S. T., Piazza, J. R., Mogle, J., Sliwinski, M. J., & Almeida, D. M. (2013). The wear and tear of daily stressors on mental health. *Psychological Science*, *24*, 733–741.

Cheng, C. (2001). Assessing coping flexibility in real-life and laboratory settings: A multimethod approach. *Journal of Personality and Social Psychology*, *80*, 814−833.

Craig, A. D. (2003). A new view of pain as a homeostatic emotion. *Trends in Neurosciences*, *26*, 303-307.

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (Eds.). (1992). *Optimal experience: Psychological studies of flow in consciousness*. Boston, MA: Cambridge university press.

Damasio, A., & Carvalho, G. B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience*, *14*, 143-152.

Davidson, R. J. (1998). Affective style and affective disorders: Perspectives from affective neuroscience. *Cognition & Emotion*, *12*, 307-330.

Davidson, R. J. (2000). Affective style, psychopathology, and resilience: Brain mechanisms and plasticity. *American Psychologist*, *55*, 1196–1214.

Davidson, R. J. (2015). Comment: Affective chronometry has come of age. *Emotion Review*, *7*, 368-370.

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. New York, NY: Sage Publishing

De Castella, K., Goldin, P., Jazaieri, H., Ziv, M., Dweck, C. S., & Gross, J. J. (2013). Beliefs about emotion: Links to emotion regulation, well-being, and psychological distress. *Basic and Applied Social Psychology, 35*, 497-505.

Desbordes, G., Gard, T., Hoge, E. A., Hölzel, B. K., Kerr, C., Lazar, S. W., ... & Vago, D. R. (2015). Moving beyond mindfulness: defining equanimity as an outcome measure in meditation and contemplative research. *Mindfulness*, *6*, 356-372.

Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, *55*, 34.

Diener, E., & Biswas-Diener, R. (2011). *Happiness: Unlocking the mysteries of psychological wealth*. New York, NY: John Wiley & Sons.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, *97*, 143-156.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121-138.

Evans, D. E., & Rothbart, M. K. (2007). Developing a model for adult temperament. *Journal of Research in Personality*, *41*, 868-888.

Feldner, M. T., Hekmat, H., Zvolensky,M. J., Vowles, K. E., Sechrist, Z., & Leen-Feldner, E.W. (2006). The role of experiential avoidance in acute pain tolerance: A laboratory test. *Journal of Behavior Therapy and Experimental Psychiatry, 37*, 146−158.

Folkman, S., & Moskowitz, J. T. (2000). Positive affect and the other side of coping. *American Psychologist*, *55*, 647.

Fredrickson, B. L., & Joiner, T. (2002). Positive emotions trigger upward spirals toward emotional well-being. *Psychological Science*, 13, 172–175.

Frijda, N. H. (2007). *The laws of emotion*. Hillsdale, NJ: Erlbaum

Frijda, N. H. (2004). Emotions and Action. In A. S. R. Manstead, N. Frijda, & A. Fischer (Eds.), *Feelings and emotions: The Amsterdam symposium.* (pp. 158–173). Boston, MA: Cambridge University Press.

Goldin PR, Gross JJ. (2010). Effects of mindfulness-based stress reduction (MBSR) on emotion regulation in social anxiety disorder. *Emotion, 10*, 83–91.

Gollwitzer, P. M., & Bargh, J. A. (Eds.). (1996). *The psychology of action: Linking cognition and motivation to behavior*. New York, NY: Guilford Press.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). New York, NY: Wiley.

Gross, J.J. (Ed.). (2015). Handbook of emotion regulation. New York, NY: Guilford.

Gross, J. J., & Feldman Barrett, L. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review*, *3*, 8-16.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348−362.

Gross, J. J., Sutton, S. K., & Ketelaar, T. (1998). Relations between affect and personality: Support for the affect-level and affective reactivity views. *Personality and Social Psychology Bulletin, 24*, 279–288.

Gruber, J., Kogan, A., Quoidbach, J., & Mauss, I. B. (2013). Happiness is best kept stable: Positive emotion variability is associated with poorer psychological health. *Emotion*, *13*, 1.

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experimental approach to behavior change.* New York, NY: Guilford Press.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change, 2nd ed*. New York, NY: Guilford Press.

Heller, A. S., van Reekum, C. M., Schaefer, S. M., Lapate, R. C., Radler, B. T., Ryff, C. D., & Davidson, R. J. (2013). Sustained striatal activity predicts eudaimonic well-being and cortisol output. *Psychological Science, 24*, 2191-2200.

Heller, A. S., Johnstone, T., Shackman, A. J., Light, S. N., Peterson, M. J., Kolden, G. G., . . . Davidson, R. J. (2009). Reduced capacity to sustain positive emotion in major depression reflects diminished maintenance of frontostriatal brain activation. *Proceedings of the National Academy of Sciences, USA*, *106*, 22445–22450.

Heller, A. S., Fox, A. S., & Davidson, R. J. (2018). Parsing affective dynamics to identify risk for mood and anxiety disorders. *Emotion*, *24*, 279–288.

Höflich, A., Michenthaler, P., Kasper, S., & Lanzenberger, R. (2018). Circuit mechanisms of reward, anhedonia, and depression. *International Journal of Neuropsychopharmacology*, *22*, 105-118.

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, *141*, 901.

Kangasniemi, A., Lappalainen, R., Kankaanpää, A., & Tammelin, T. (2014). Mindfulness skills, psychological flexibility, and psychological symptoms among physically less active and active adults. *Mental Health and Physical Activity*, *7*(3), 121-127.

Kanske, P., & Kotz, S. A. (2012). Effortful control, depression, and anxiety correlate with the influence of emotion on executive attentional control. *Biological Psychology*, *91*, 88-95.

Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*, *30*, 865-878.v

Kaviani, H., Gray, J. A., Checkley, S. A., Raven, P. W., Wilson, G. D., & Kumari, V. (2004). Affective modulation of the startle response in depression: influence of the severity of depression, anhedonia, and anxiety. *Journal of affective disorders*, *83*(1), 21-31.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*, 932-932.

Koval, P., Sütterlin, S., & Kuppens, P. (2016). Emotional inertia is associated with lower well-being when controlling for differences in emotional context. *Frontiers in Psychology*, *6*, 1997.

Kreft, I. G., & De Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

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

Lang, P. J. (1995). The emotion probe: studies of motivation and attention. *American Psychologist*, *50*, 372.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, *1*, 39-58

Larsen, R. J., Augustine, A. A., & Prizmic, Z. (2009). A process approach to emotion and personality: Using time as a facet of data. *Cognition and Emotion*, *23*, 1407-1426.

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, *61*, 132–140

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: Physiological linkage and affective exchange. *Journal of Personality and Social Psychology, 45,* 587-597.

Low, C. A., Stanton, A. L., & Bower, J. E. (2008). Effects of acceptance-oriented versus evaluative emotional processing on heart rate recovery and habituation. *Emotion*, *8*, 419.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. Boston, MA: MIT press.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131*, 803–855.

Lyubomirsky, S., & Nolen-Hoeksema, S. (1993). Self-perpetuating properties of dysphoric rumination. *Journal of Personality and Social Psychology*, *65*, 339.

Masuda, A., & Tully, E. C. (2012). The role of mindfulness and psychological flexibility in somatization, depression, anxiety, and general psychological distress in a nonclinical college sample. *Journal of Evidence-Based Complementary & Alternative Medicine*, *17*(1), 66-71.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, *5*, 175.

Mikolajczak, M., Quoidbach, J., Vanootighem, V., Lambert, F., Lahaye, M., Fillée, C., & de Timary, P. (2010). Cortisol awakening response (CAR)'s flexibility leads to larger and more consistent associations with psychological factors than CAR magnitude. *Psychoneuroendocrinology, 35*, 752–757.

Miller, A. L., Rathus, J. H., & Linehan, M. M. (2006). *Dialectical behavior therapy with suicidal adolescents*. New York, NY: Guilford Press.

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review, 5*(2), 119-124.

Nesse, R. M., & Ellsworth, P. C. (2009). Evolution, emotions, and emotional disorders. *American Psychologist*, *64*, 129.

Nezlek, J. B. (2001). Multilevel random coefficient analyses of event- and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin*, *27*, 771–785.

Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass, 2*, 842-860.

Oishi, S., Diener, E., Choi, D.-W., Kim-Prieto, C., & Choi, I. (2007). The dynamics of daily events and well-being across cultures: When less is more. *Journal of Personality and Social Psychology*, *93,* 685–698.

Ong, A. D., Bergeman, C. S., Bisconti, T. L., & Wallace, K. A. (2006). Psychological resilience, positive emotions, and successful adaptation to stress in later life. *Journal of Personality and Social Psychology*, *91*(4), 730–749.

Peterson, C., & Seligman, M. E. (1983). Learned helplessness and victimization. *Journal of Social Issues*, *39*, 103-116.

Robinson, D. M., Klein, R. J. (2018). What do subjective well-being judgments mean? Sources & distinctions, processes & mechanisms. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being.* Salt Lake City, UT: DEF Publishers.

Robinson, M. D., & Tamir, M. (2005). Neuroticism as mental noise: a relation between neuroticism and reaction time standard deviations. *Journal of Personality and Social Psychology*, *89*, 107.

Rueschkamp, G., Johanna, M., Kuppens, P., Riediger, M., Blanke, E. S., & Brose, A. (2018). Higher well-being is related to reduced affective reactivity to positive events in daily life. *Emotion, 33, 4*85–498.

Sahdra, B. K., Shaver, P. R., & Brown, K. W. (2010). A scale to measure nonattachment: A Buddhist complement to Western research on attachment and adaptive functioning. *Journal of Personality Assessment*, *92*, 116-127.

Sahdra, B., Ciarrochi, J., & Parker, P. (2016). Nonattachment and mindfulness: Related but distinct constructs. *Psychological Assessment*, *28*, 819.

Sapolsky, R. M. (2004). *Why zebras don't get ulcers: The acclaimed guide to stress, stress-related diseases, and coping-now revised and updated*. New York, NY: Holt paperbacks.

Schaefer, S. M., Boylan, J. M., Van Reekum, C. M., Lapate, R. C., Norris, C. J., Ryff, C. D., & Davidson, R. J. (2013). Purpose in life predicts better emotional recovery from negative stimuli. *PloS one, 8*, e80329.

Scherer, K. R. (2000). Emotions as episodes of subsystems synchronization driven by nonlinear appraisal processes. In M. D. Lewis & I. Granic (Eds.), *Emotion, development, and self-organization: Dynamic systems approaches to emotional development.* (pp. 70–99). New York, NY: Cambridge University Press.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, *23*, 323-355.

Taubitz, L. E., Robinson, J. S., & Larson, C. L. (2013). Modulation of the startle reflex across time by unpleasant pictures distinguishes dysphoric from non-dysphoric women. *International Journal of Psychophysiology*, *87*, 124-129.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software, 7*.

Trull, T. J., Lane, S. P., Koval, P., & Ebner-Priemer, U. W. (2015). Affective dynamics in psychopathology. *Emotion Review, 7,* 355-361.

Tugade, M. M., & Fredrickson, B. L. (2004). Resilient individuals use positive emotions to bounce back from negative emotional experiences. *Journal of Personality and Social Psychology*, *86*, 320.

Tull, M. T., & Roemer, L. (2007). Emotion regulation difficulties associated with the experience of uncued panic attacks: Evidence of experiential avoidance, emotional nonacceptance, and decreased emotional clarity. *Behavior Therapy*, *38*, 378-391.

van de Leemput, I. A., Wichers, M., Cramer, A. O., Borsboom, D., Tuerlinckx, F., Kuppens, P., ... & Derom, C. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences, 111,* 87-92.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97,* 346–353.

Waugh, C. E., Thompson, R. J., & Gotlib, I. H. (2011). Flexible emotional responsiveness in trait resilience. *Emotion*, *11*, 1059-1077

West, S. G., Ryu, E., Kwok, O., & Cham, H. (2011). Multilevel modeling: Current and future applications in personality research. *Journal of Personality, 70*, 2-50.

Wichers, M., Wigman, J. T. W., & Myin-Germeys, I. (2015). Micro-level affect dynamics in psychopathology viewed from complex dynamical system theory. *Emotion Review, 7,* 362-367.

Wichers, M., Groot, P. C., Psychosystems, E. S. M., & EWS Group. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, *85*, 114-133.

Wichers, M., Myin-Germeys, I., Jacobs, N., Peeters, F., Kenis, G., Derom, C., . . .van Os, J. (2007). Genetic risk of depression and stress-induced negative affect in daily life. *The British Journal of Psychiatry: The Journal of Mental Science*, *191*, 218–223.

Wilkowski, B. M., Robinson, M. D., & Troop-Gordon, W. (2010). How does cognitive control reduce anger and aggression? The role of conflict monitoring and forgiveness processes. *Journal of Personality and Social Psychology, 98*, 830-840.

Zelenski, J. M., & Larsen, R. J. (1999). Susceptibility to affect: A comparison of three personality taxonomies. *Journal of Personality*, *67,* 761–791.

APPENDIX A. FLOURISHING SCALE

Below are 8 statements with which you may agree or disagree. Using the 1–7 scale below, indicate your agreement with each item by indicating that response for each statement.

7 - Strongly agree
6 - Agree
5 - Slightly agree
4 - Neither agree nor disagree
3 - Slightly disagree
2 - Disagree
1 - Strongly disagree

____ I lead a purposeful and meaningful life
____ My social relationships are supportive and rewarding
____ I am engaged and interested in my daily activities
____ I actively contribute to the happiness and well-being of others
____ I am competent and capable in the activities that are important to me
____ I am a good person and live a good life
____ I am optimistic about my future
____ People respect me

Scoring:
Average all items to create a single score for each participant.

APPENDIX B. SCALE OF POSITIVE AND NEGATIVE EXPERIENCE (SPANE)

Please think about what you have been doing and experiencing during the past four weeks. Then report how much you experienced each of the following feelings, using the scale below. For each item, select a number from 1 to 5.

1. Very Rarely or Never
2. Rarely
3. Sometimes
4. Often
5. Very Often or Always

____Positive
____Negative
____Good
____Bad
____Pleasant
____Unpleasant
____Happy
____Sad
____Afraid
____Joyful
____Angry
____Contented

Scoring
The measure can be used to derive an overall affect balance score, but can also be divided into positive and negative feelings scales.

Positive Emotionality (PE). Add the scores, varying from 1 to 5, for the six items: positive, good, pleasant, happy, joyful, and contented.

Negative Emotionality (NE). Add the scores, varying from 1 to 5, for the six items: negative, bad, unpleasant, sad, afraid, and angry.

APPENDIX C. DAILY DIARY SURVEY

Daily Positive Emotions
To what extent did you feel each of the following today? Use the scale below.

1. Not at all
2.
3.
4.
5. Extremely

____Today, I felt happy
____Today, I felt positive
____Today, I felt excited

Daily Negative Emotions
To what extent did you feel each of the following today? Use the scale below.

6. not at all
7.
8.
9.
10. extremely

____Today, I felt sad
____Today, I felt negative
____Today, I felt distressed

Daily Satisfaction
To what extent did you feel each of the following today? Use the scale below.

11. not at all
12.
13.
14.
15. extremely

____Today, I was satisfied with myself
____Today, I was satisfied with my health