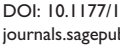
Assessment
2020, Vol. 27(8) 1928–1945

DOI: 10.1177/1073191119839138
journals.sagepub.com/home/asm


Development of a Novel Method of Emotion Differentiation That Uses Open-Ended Descriptions of Momentary Affective States

Charlotte Ottenstein[1] **and Tanja Lischetzke**[1]

Abstract

Emotion differentiation (ED) has been conceptualized as a trait that facilitates emotion regulation and increases well-being. Yet basic questions remain unanswered about how best to assess it and whether favorable outcomes can be observed only during times of stress. The goal of the present research was to develop a novel behavioral (specificity) index of ED. We conducted two daily diary studies (N = 111-190) in which we included different measures of ED, well-being, and emotion regulation. The different ED measures were largely unrelated to each other. In both studies, the specificity index of ED showed a positive association with daily well-being, but in Study 2, this association held only on days with a negative event. Results regarding ED and the use of emotion-regulation strategies were inconsistent across strategies and studies. Possible reasons for these mixed results (e.g., sample selection, context sensitivity of regulation strategies) are discussed.

Keywords

emotion differentiation, emotional granularity, well-being, emotion regulation, daily diary

In every language, many words can be used to describe how someone feels at a certain moment. Yet not everyone makes use of this rich emotion vocabulary in everyday life when talking about his or her feelings. Whereas some individuals tend to use specific and discrete emotion terms (e.g., "sad" or "frustrated") to describe an emotional experience, others tend to use broader terms (e.g., "bad" or "negative") that mainly communicate displeasure or pleasure. These individual differences have been called emotion differentiation (ED; Feldman Barrett, Gross, Christensen, & Benvenuto, 2001) or emotional granularity (e.g., Tugade, Fredrickson, & Feldman Barrett, 2004). According to these theoretical approaches, individuals with high ED are able to generate a distinctive, granular, and precise representation of their emotional experience, whereas individuals with low ED use terms that are located along a single dimension of pleasantness-unpleasantness (Feldman Barrett, 1998; Feldman Barrett et al., 2001). In the present article, we aim to present a novel method for assessing ED that more closely reflects this definition of the construct than other measures that have been previously proposed. Before we discuss how ED can be measured and present the rationale behind our novel method, we summarize theoretical considerations about ED's role in the emotion-regulation process and its relation to subjective well-being. These theoretical considerations will be relevant for deducting hypotheses on the predictive validity of methods assessing ED.

The Functionality of Emotion Differentiation for Emotion Regulation and Well-Being

As Kashdan, Feldman Barrett, and McKnight (2015) have argued, the use of specific, differentiated emotion words conveys important information about an emotion-eliciting event. According to this information, individuals who label their emotional experiences with specific terms can regulate their intense negative emotions more effectively, are better able to pursue personal strivings, and finally, are able to achieve greater well-being. These assumptions are in line with the feelings-as-information theory (Schwarz, 1990), which posits that people use emotions as a source of information about themselves or about their environment and

[1]University of Koblenz-Landau, Landau, Germany

Corresponding Author:
Charlotte Ottenstein, Center for Methodologies, Diagnostics and Evaluation, University of Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany.
Email: ottenstein@uni-landau.de

that different types of feelings provide different types of information (Schwarz, 2012). Hence, due to the higher specificity with which they can identify their affective states, individuals with high ED should have access to more detailed and differentiated information and, thus, should be able to show more adaptive and adequate responses to experienced events (Erbas, Ceulemans, Koval, & Kuppens, 2015; Feldman Barrett et al., 2001; Kirby, Tugade, Morrow, Ahrens, & Smith, 2014).

In his extended process model of emotion regulation, Gross (2015) combined an account of the emotion-generating process (via appraisals) on a lower level with an upper level emotion-regulation process, which involves (a) identifying the emotions to be regulated, (b) selecting an emotion-regulation strategy, and (c) implementing a regulation strategy. This model helps pinpoint more clearly the steps in the emotion-regulation process at which higher ED should provide benefits (cf. Lischetzke & Eid, 2017): During the identification stage, higher ED should help individuals form a specific intention to regulate their feelings (i.e., upregulate, downregulate, or maintain) that is in accordance with their current goals and circumstances. Once individuals have formed an intention to regulate, they can select and implement a regulation strategy. The more specific individuals' representations of their affective states are, the better individuals may be at selecting a strategy (e.g., cognitive change vs. distraction) that has proven successful in similar situations in the past and that has the best chance of regulating the affective state in accordance with their current goals.

Most empirical research on ED to date has focused on the differentiation of *negative* emotions (i.e., negative ED), and fewer studies have investigated the differentiation of positive emotions—possibly because deficits in negative ED (as compared with deficits in positive ED) will have more severe immediate effects on emotion regulation (e.g., Feldman Barrett et al., 2001). In the present research, therefore, our focus is on negative ED.

Previous studies have partially supported the idea that negative ED is beneficial for affect regulation and well-being by showing that higher negative ED was related to more adaptive emotion regulation (Feldman Barrett et al., 2001), less alcohol use in response to intense negative emotions (Kashdan, Ferssizidis, Collins, & Muraven, 2010), lower aggressiveness in response to provocation (Pond et al., 2012), and higher self-esteem (Erbas, Ceulemans, Pe, Koval, & Kuppens, 2014). In clinical studies, negative ED was lower in individuals with major depressive disorders (Demiralp et al., 2012), schizophrenia (Kimhy et al., 2014), social anxiety disorder (Kashdan & Farmer, 2014), and borderline personality disorder (Zaki, Coifman, Rafaeli, Berenson, & Downey, 2013) than in healthy adults. Other studies, however, have reported no association between negative ED and well-being indicators (life satisfaction, trait positive affect, trait negative affect, resilience; Grühn, Lumley, Diehl, & Labouvie-Vief, 2013; Kirby et al., 2014), or inconsistent results regarding negative ED's role in emotion regulation (Hay & Diehl, 2011). For example, high negative ED was related to more reappraisal use—but unrelated to suppression use—in individuals with low social anxiety (O'Toole, Jensen, Fentz, Zachariae, & Hougaard, 2014).

One explanation for these mixed findings with regard to whether negative ED is adaptive or not might be that high ED is most helpful when intense emotions are experienced. Feldman Barrett et al. (2001) hypothesized that the relation between ED and regulation would be "strongest in the context of intense negative emotions, where the press for emotion regulation is greatest" (p. 715). They found that negative ED was related to greater emotion regulation, and this relation was more pronounced for individuals who experienced relatively intense emotions. In an ecological momentary assessment study on alcohol intake (Kashdan et al., 2010), higher negative ED predicted lower alcohol consumption during drinking episodes only in situations where predrinking negative affect was high but not in situations where predrinking negative affect was low. Similarly, Pond et al. (2012) found an interaction between negative ED and anger intensity in the prediction of aggressive tendencies: The relation between higher negative ED and lower aggressive tendencies was more pronounced on days when participants felt more anger. Taken together, there is some theoretical and empirical evidence that negative ED might be adaptive under certain circumstances. Nonetheless, the question of the most suitable operationalization of ED has yet to be fully resolved.

Measuring Emotion Differentiation

To date, there are two main groups of ED measures: self-report measures on one hand and indices that are based on repeated affect ratings on the other hand. One self-report measure of ED is the differentiation subscale from the Range and Differentiation of Emotional Experience Scale (RDEES; Kang & Shaver, 2004). Kang and Shaver's (2004) validation studies showed that the differentiation subscale had positive relations with emotional clarity and emotional awareness and a negative association with alexithymia. One problem with global self-report measures of emotion-related traits, however, is that they tend to capture individuals' beliefs about themselves rather than provide an accurate representation of momentary experience (e.g., Robinson & Clore, 2002). Moreover, because ED is considered a skill, researchers have argued that it should be measured behaviorally (Kashdan et al., 2015).

The second group of ED measures, which can be considered behavioral (performance-based) measures, uses repeated measures of intensity ratings of different affect terms (see, Trull, Lane, Koval, & Ebner-Priemer, 2015).

The basic rationale is that individuals who give similar ratings to the emotion terms over time (resulting in a high covariation) have lower ED (Feldman Barrett, 1998; Feldman Barrett et al., 2001). Two different performance-based indices have been used: average Pearson correlations (e.g., Demiralp et al., 2012; Feldman Barrett, 1998; Feldman Barrett et al., 2001; O'Toole et al., 2014) and intraclass correlation coefficients (ICCs; e.g., Boden, Thompson, Dizén, Berenbaum, & Baker, 2013; Erbas, Ceulemans, Boonen, Noens, & Kuppens, 2013; Grühn et al., 2013; Kashdan & Farmer, 2014; Pond et al., 2012; Tugade et al., 2004). The two coefficients are calculated separately for each participant. To calculate a participant's average Pearson correlation, the correlations between all possible pairs of emotion items are determined, and after applying the Fisher *r*-to-*z* transformation, the mean correlation is computed (Feldman Barrett, 1998). Low ED is characterized by a high correlation, and high ED is characterized by a low correlation between emotion terms over time. For ease of interpretation, the average correlations are typically multiplied by −1 or subtracted from 1 so that high values reflect high ED. ICCs were originally proposed in the field of interrater reliability (Shrout & Fleiss, 1979) where *n* targets are rated by *k* judges. Transferred to the field of ED, *k* emotion terms are rated at each of the *n* measurement occasions for each participant. High correlations between different emotions result in high ICCs, which are interpreted to indicate low ED. Again, most researchers reverse these indices for ease of interpretation.

Specific problems are related to these performance-based indices of ED. First, individuals might differ in the degree to which the events they encounter during the daily diary period are similar (vs. diverse) and elicit similar (vs. diverse) emotions. As Erbas et al. (2014) noted, these differences might be "partly responsible for differences in emotion differentiation. For instance, some participants may go through a very calm period in life, whereas others may go through a more demanding period, with consequences for the range of emotions potentially experienced" (p. 1202). As a solution to this problem, ED has been measured in emotion-eliciting situations in controlled laboratory settings (e.g., Erbas et al., 2013; Erbas et al., 2015) or using hypothetical scenarios (e.g., Boden et al., 2013; Kirby et al., 2014). This approach ensures that the rated emotions all refer to the same emotion source.

Another solution to the problem of individual differences in situational diversity over time might be a state perspective on ED: Recently, Tomko et al. (2015) suggested that researchers should quantify negative ED by using *momentary* intraclass correlation coefficients, which indicate the consistency with which negative emotion items are rated across emotion-specific subscales (e.g., fear, hostility, and sadness) at each measurement occasion. Plonsker, Gavish Biran, Zvielli, and Bernstein (2016) assessed emotional experience in response to images presented to participants in a laboratory session. As an index of (state) ED, they calculated the degree of co-activation of multiple negative emotions (e.g., intensity of disgust, anger, sadness, embarrassment, shame, guilt, or anxiety) when a specific state (e.g., fear) was elicited. A higher degree of co-activation was used as an index of low ED. In our view, however, the rationale behind the state ED measures applied by Tomko et al. (2015) and Plonsker et al. (2016) has a theoretical drawback. Using the consistency of emotion ratings across items/subscales as an index of ED makes it difficult to distinguish between two kinds of situations: (a) situations in which multiple emotions of the same valence are actually elicited and represented as such in the individual versus (b) situations in which one specific emotion is elicited but the individual's subjective experience is a "muddle," and hence, different emotion terms are rated in a similar way. In our view, ideally, a measure of emotion differentiation should not by definition equate the co-occurrence of multiple emotions of the same valence (or of different valences) with a lack of granularity in how individuals experience their affective state (cf. Grossmann, Huynh, & Ellsworth, 2016).

Finally, a very fundamental problem with measures of ED that are based on (repeated) affect ratings (both in daily life and in the laboratory) is that they rely on individuals' responses to a list of emotion terms provided by the researchers. Kashdan et al. (2015) argued that "to truly capture an individual's spontaneous ED performance, researchers must assess what is being felt without using prompts with a closed-ended list of emotion-word labels" (p. 12), and as an alternative, they recommended a research strategy in which "individuals verbalize what they are feeling while engaged in a situation" (p. 12).

In the present research, we followed this suggestion and developed a new ED index that is based on an open-response format instead of a closed-ended list of emotions. In general, the suggested approach can be used to assess negative ED (by investigating negative emotional experiences) or positive ED (by investigating positive emotional experiences). In the present article, our focus was on negative ED. Therefore, we describe the rationale of the novel method with respect to the differentiation of negative emotions. We operationalized negative ED in a way that closely reflects the definition of the construct as proposed by Feldman Barrett and colleagues (e.g., Feldman Barrett, 1998; Feldman Barrett et al., 2001). Feldman Barrett (2006) stated:

> When people report on their experiences [. . .], and those self-reports are analyzed as verbal behaviors, it becomes clear that some people make categorical distinctions, characterizing their experiences in discrete emotion terms, whereas others characterize their experiences in broad, global terms. (p. 24)

To directly operationalize this construct definition, the rationale behind the new ED index is the following: Using

a daily diary approach, participants are repeatedly asked to report on daily negative events and their affective experiences during these events. Instead of rating their affective state on a provided list of emotion terms, participants are asked to describe their affective state with terms (adjectives) that they select themselves. That is, as suggested by Kashdan et al. (2015), participants are given the opportunity to choose their own words when asked to describe their emotional experiences. After data collection, the terms that participants used to describe their affective states are coded as either *specific* (indicating high negative ED; e.g., scared, angry, sad) or *general* (indicating low negative ED; e.g., bad, negative, unpleasant). An individual's negative ED score is operationalized as the proportion of specific adjectives out of all adjectives that were used to describe momentary affect in response to negative events (across all measurement occasions). We use the term *specificity index of negative ED* to denote the new assessment method.

Aims of the Present Research

Our major aim was to develop and test the novel specificity index of negative ED that is based on repeated assessments of affective states in an open-response format and uses a coding scheme for specific versus general affect terms. To gain insight into the properties of this novel specificity index of negative ED, we conducted two daily diary studies and aimed to scrutinize (a) characteristics of the response process (providing general information on individuals' use of the open-response format), (b) rater agreement on specificity codings (as an index of the method's objectivity/reliability), (c) convergent relations with other ED measures (global self-report in Studies 1 and 2, ICC measure in Study 2), (d) relations between the specificity index of negative ED and potential confounding variables (verbal ability in Study 2), and (e) associations with measures of affect regulation and well-being as an indication of predictive validity.

Study 1

Study 1 was the first study to use the novel specificity index of negative ED. Hence, we sought to gain some insight into individuals' repeated use of the open-response format. In particular, we wanted to explore whether momentary time pressure has an influence on the number of adjectives reported, and we wanted to check (similar to a manipulation check) the valence of the adjectives that were reported as responses to negative events.

To our knowledge, only two studies to date (unpublished raw data from Feldman Barrett, 2001, cited in Lindquist & Feldman Barrett, 2008; Lindquist, Gendron, Oosterwijk, & Feldman Barrett, 2013) have assessed both performance-based indicators and self-reports of ED in the same sample. Therefore, we decided to use a multimethod assessment strategy and to additionally gather self-reports of ED to be able to analyze the convergence between the measures and compare the link between ED and the criterion variables across assessment methods.

A construct that is also concerned with the precise representation of affective experiences is alexithymia. Alexithymia was first described by Sifneos (1973) and literally means "no words for emotions." Two components of alexithymia, namely, difficulty identifying one's own feelings and distinguishing between feelings and bodily sensations, and difficulty describing one's feelings to others, (e.g., Luminet, Rimé, Bagby, & Taylor, 2004) refer to an "impoverished conceptual system for emotion and emotion vocabulary, associated with impoverished descriptions of emotional experiences" (Kashdan et al., 2015, p. 12) and, hence, should be the ones that are more closely related to the construct of ED. Supporting this assumption, Erbas et al. (2014) found that negative ED (as assessed with the ICC measure) had low to moderate correlations with less difficulty identifying feelings (Studies 2 and 3) and less difficulty describing feelings (in Study 3 but not in Study 2). However, Boden et al. (2013) found no relation between clarity of feelings (which can be considered the opposite pole of difficulty identifying feelings) and negative ED (as assessed with the ICC measure).

On the basis of theoretical accounts of the functionality of ED for well-being (Kashdan et al., 2015), we expected positive relations between negative ED and well-being indicators. To extend previous research, we aimed to test whether the adaptive value of negative ED would hold for both the affective and cognitive facets of well-being (Diener, Suh, Lucas, & Smith, 1999). We sought to measure well-being in an ecologically valid way by using state measures of affective and cognitive well-being in daily life (momentary pleasant–unpleasant mood and daily life satisfaction).

As already outlined in the introduction, it might be the case that high negative ED is most predictive of well-being in situations that entail intense negative emotions (e.g., Feldman Barrett et al., 2001; Kashdan et al., 2010; Pond et al., 2012). We tested the hypothesis of negative ED's differential predictive validity in the following way: We scrutinized whether negative ED was more closely related to daily well-being on days that presented a challenge to individuals' well-being (i.e., days with a negative event) compared with less challenging days (i.e., days without a negative event). In particular, we expected that negative ED would buffer the detrimental effect of negative events on well-being.

On the basis of broad theoretical considerations of ED's role in the emotion-regulation process (see e.g., Gross, 2015; Kashdan et al., 2015; Lischetzke & Eid, 2017), we expected that higher negative ED would be generally related to more adaptive emotion regulation. However, the specific paths through which higher negative ED might translate

into successful emotion regulation are not clear, and there is a lack of empirical studies on the link between negative ED and emotion regulation. In the present research, we focused on two well-studied emotion-regulation strategies—cognitive reappraisal (i.e., construing a potentially emotion-eliciting situation in a way that changes its emotional impact) and expressive suppression (i.e., inhibiting ongoing emotion-expressive behavior; Gross & John, 2003).

Reappraisal can be considered an effective strategy in downregulating negative emotions (e.g., Gross, 1998). According to Boden and Thompson (2015), a mental representation of the emotion is needed when reappraisal is used. Hence, higher negative ED should facilitate the use of reappraisal. Suppression, on the other hand, might be considered a rather maladaptive strategy because it has been found to increase negative affect (e.g., Brans, Koval, Verduyn, Lim, & Kuppens, 2013; Geisler & Schröder-Abé, 2015) and decrease positive affect (e.g., Brans et al., 2013), and it has been found to be related to higher depressiveness, lower self-esteem, and lower well-being (Gross & John, 2003). However, other findings have shown that suppression is unrelated to emotional intensity (e.g., Gross, 1998) or even helps in the downregulation of anger (Germain & Kangas, 2015). Therefore, the classification of suppression as a maladaptive strategy is not as straightforward as it may seem. Irrespective of the functionality of suppression, ED and suppression might be associated: Individuals using expressive suppression direct their attention "away from the emotion source and/or response" (Boden & Thompson, 2015, p. 401), a process that might interfere with the specific identification of emotion (Gross & John, 2003). In their process model of emotional expression, Kennedy-Moore and Watson (1999) proposed that being unclear about one's feelings interferes with emotional expression, leading individuals to inhibit expressive behavior (i.e., to suppress it). From both perspectives, one might conclude that higher negative ED should be associated with a less frequent use of suppression.

To our knowledge, there is only one study that analyzed the associations of ED with reappraisal and suppression: O'Toole et al. (2014) found that high positive ED was related to less use of suppression and that high negative ED was related to more reappraisal use (in individuals with low social anxiety). As a first step in testing our hypotheses, in Study 1, we used established trait measures of the two emotion-regulation strategies (Gross & John, 2003).

Method

Participants. Participants were recruited via e-mail and posts on German online forums (with different topics, e.g., politics, health, education, cooking). They participated voluntarily and had the chance to win one of several 30€ vouchers for online shopping. Because this study was the first study to examine the novel specificity index of ED and we could not estimate a precise effect size beforehand, we aimed to recruit at least 100 participants. Such a sample size would, for example, allow us to test a correlation of $r = .25$ with a power of .80 at a significance level of $\alpha = .05$ (one-sided). A total of 115 individuals from the general population filled out the initial online questionnaire and took part in the daily diary phase that followed. Only participants with at least one reported positive and one reported negative event were included in the present analyses. Therefore, we excluded data from four participants who did not report any negative events during the daily diary phase. The final sample consisted of 111 participants; 73% were female, and the mean age of the sample was 34.95 years ($SD = 15.72$). Participants' highest educational achievement was distributed as follows: 19% had no professional achievement or were in training, 14% had completed an apprenticeship, 19% had a polytechnical degree, and 48% had a university degree.

Procedure. The study began with an online questionnaire to assess trait measures and sociodemographic information. Subsequently, a 3-week daily diary phase followed (via an online questionnaire). Each evening at 6 p.m., participants received an e-mail with a link. Daily measures had to be filled out between 6 p.m. and 2 a.m. Participants had to report their current mood state, satisfaction with the day, as well as one negative event they experienced during the day and their related feelings.[1] Compliance was good: Participants completed $M = 15.92$ ($SD = 5.07$) of the 21 questionnaires (range: 1-21). The number of negative events that were reported ranged from 1 to 19 ($M = 8.23$, $SD = 4.64$).

Measures

Emotion differentiation–specificity index. Participants were asked to describe their affective state during the reported events with an open-response format ("Please use adjectives to describe how you felt during this event"). They were not given a list of emotion terms but were asked to type their own words (adjectives) into text boxes. These terms were analyzed via the text analysis software RIOT Scan (Boyd, 2014). We used an existing dictionary of emotion words (Linguistic Inquiry and Word Count [LIWC]; Pennebaker, Francis, & Booth, 2001; German version by Wolf et al., 2008) as the software's underlying word pool and added any terms mentioned by the participants that were not already part of the LIWC dictionary. All terms were coded by two independent raters as either *specific* (indicating high ED; e.g., scared, angry, sad) or *general* (indicating low ED; e.g., bad, negative, unpleasant). A somewhat related distinction of labels of emotions was made in the study by Labouvie-Vief, DeVoe, and Bulka (1989). The authors evaluated participants' spontaneous descriptions of emotional experience during a recent emotion-eliciting situation. Among other

aspects, the psychological labels of emotions were coded as relatively undifferentiated and static or as highly differentiated and dynamic. Our distinction between specific and general affective terms also bears similarity to a distinction of emotion reports made in the Levels of Emotional Awareness Scale (LEAS; Lane, Quinlan, Schwartz, Walker, & Zeitlin, 1990). LEAS is a performance measure in which individuals' verbal descriptions of hypothetical reactions to vignettes are scored by raters. Our categories of general affective adjectives versus specific affective adjectives corresponded to the LEAS categories of undifferentiated emotions (LEAS Level 2) and differentiated emotions (LEAS Level 3), respectively. LEAS Level 4 refers to the use of two or more Level-3 words. In our specificity index of ED, information about the frequency of specific (i.e., differentiated) words was indirectly included because we calculated the proportion of specific adjectives out of all affective adjectives that were mentioned over time. Adjectives that did not describe a feeling state (e.g., old, drunken, hungry) were not considered for the specificity index of ED.

To obtain information about the objectivity and reliability of the specificity ratings, we calculated Cohen's kappa as a measure of interrater agreement. The size of Cohen's kappa was good (Cohen's $\kappa = .62$, $N_{words} = 1{,}082$). The two raters discussed the words that they had rated differently and agreed on a common solution. The final dictionary consisted of 727 specific and 62 general adjectives.[2] This coding system served as the basis for the specificity index of ED. In particular, we used the terms provided by participants in response to negative events to calculate a negative ED score. For each participant, the total number of specific adjectives was divided by the total number of specific plus general adjectives. That is, we used the proportion of specific adjectives as an ED indicator.

Emotion differentiation (global self-report). We used the differentiation subscale from the RDEES (Kang & Shaver, 2004). This subscale consists of seven items (e.g., "Each emotion has a very distinct and unique meaning to me"), which were answered on a 4-point response scale (1 = *strongly disagree* to 4 = *strongly agree*). Cronbach's α was .86.

Alexithymia. Alexithymia facets were assessed with two subscales from the Toronto Alexithymia Scale (TAS-20; Bagby, Parker, & Taylor, 1994), namely, *Difficulty Identifying Feelings* (DIF; seven items) and *Difficulty Describing Feelings* (DDF; five items). Each item was answered on a 4-point scale (1 = *strongly disagree* to 4 = *strongly agree*). Cronbach's α was .81 for DIF and .72 for DDF.

Reappraisal and suppression (global self-report). The emotion-regulation strategies reappraisal and suppression were assessed with the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003). Reappraisal was assessed with six items (e.g., "I control my emotions by changing the way I think about the situation I'm in"), and suppression was assessed with four items (e.g., "I control my emotions by not expressing them"), using a 7-point response scale (1 = *strongly disagree* to 7 = *strongly agree*). Cronbach's α was .86 for reappraisal and .62 for suppression.

Daily life satisfaction. Each evening, we assessed daily life satisfaction with two items from Oishi, Diener, Choi, Kim-Prieto, and Choi (2007): "How was today?" (1 = *terrible* to 4 = *excellent*) and "How satisfied were you today?" (1 = *very dissatisfied* to 4 = *very satisfied*). For those daily measures that were assessed with only two items, we estimated local (within-occasion) reliability (Buse & Pawlik, 1996) by calculating the polychoric correlation between the items for each day and summarizing them by identifying the median. The median polychoric correlation across measurement occasions was .93. Local reliability indicates the internal consistency of the measure at the same occasion, whereas aggregate reliability indicates the consistency of aggregate scores across occasions. To estimate aggregate reliability, we calculated the Pearson correlation between the two items (aggregated across occasions), which was .93.

Momentary pleasant–unpleasant mood. Each evening, participants rated their momentary mood on an adapted short version of the Multidimensional Mood Questionnaire (Steyer, Schwenkmezger, Notz, & Eid, 1994), which was optimized for use in a diary study (smaller number of bipolar items instead of the original monopolar format; Lischetzke, Pfeifer, Crayen, & Eid, 2012). We used three items to assess pleasant–unpleasant mood (unwell–well, bad–good, unhappy–happy). Participants rated how they felt at the moment on 6-point bipolar intensity scales (e.g., 1 = *very unhappy* to 6 = *very happy*). The within-person–level α (Geldhof, Preacher, & Zyphur, 2014) was .90, and the between-persons–level α was .97.

Momentary time pressure. At each measurement occasion, participants rated their momentary time pressure with a single item ("I am currently pressed for time") on a 4-point response scale (1 = *strongly disagree* to 4 = *strongly agree;* $M = 1.84$, $SD = 0.64$).

Data Analysis. Relations between variables on the person level were analyzed with correlations.[3] To predict individual differences in the constructs that were assessed on a daily basis (life satisfaction and pleasant–unpleasant mood), multilevel models were created to account for the nested data structure (days nested in individuals). It is important to note that daily well-being was assessed each day no matter whether an individual had reported experiencing a positive or a negative event that day.

The equations for predicting the well-being (daily life satisfaction or momentary pleasant–unpleasant mood) of person *i* on day *t* with a dummy variable indicating the

Table 1. Means, Standard Deviations, and Bivariate Correlations for Main Variables (Study 1).

Variable	1	2	3	4	5	6	7	8
Emotion differentiation								
1. $ED_{self\text{-}report}$	—							
2. Negative $ED_{specificity}$	.12	—						
Alexithymia								
3. DIF	−.31***	.02	—					
4. DDF	−.48***	−.08	.61***	—				
Emotion regulation								
5. Reappraisal	.22*	.03	−.06	−.01	—			
6. Suppression	−.19*	−.08	.24**	.37***	.08	—		
Daily well-being								
7. Daily LS	.09	.26*	−.30**	−.20*	.23*	−.14	—	.65***
8. P-U mood	.04	.26*	−.30**	−.17†	.27**	−.20*	.89***	—
M	2.82	0.86	1.88	2.01	4.38	3.23	2.95	4.52
SD	0.54	0.22	0.55	0.56	1.11	1.06	0.37	0.70

Note. Between-person correlations ($N_{Persons}$ = 111) are presented below the diagonal; the within-person correlation between the two daily well-being measures (N_{Days} = 1,767) is presented above the diagonal. $ED_{self\text{-}report}$ = Global self-report of emotion differentiation; Negative $ED_{specificity}$ = specificity index of negative emotion differentiation; DIF = difficulty identifying feelings; DDF = difficulty describing feelings; Daily LS = daily life satisfaction; P-U mood = momentary pleasant–unpleasant mood.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

presence of a negative event (0/*day without a negative event*, 1/*day with a negative event*) and (grand-mean centered) person-level negative ED were

$$\text{Level 1: Well-being}_{ti} = \pi_{0i} + \pi_{1i} \cdot \text{Neg.event}_{ti} + e_{ti} \quad (1),$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + \beta_{01} \cdot \text{Neg.ED}_i + r_{0i} \quad (2a),$$

$$\pi_{1i} = \beta_{10} + \beta_{11} \cdot \text{Neg.ED}_i + r_{1i} \quad (2b).$$

The fixed effects in Equations (2a) and (2b) have the following meaning: β_{00} represents mean well-being on days without a negative event, β_{01} represents the expected change in well-being for a 1-unit change in negative ED (i.e., negative ED slope) on days without a negative event, β_{10} represents the mean difference in well-being between days with versus without a negative event (for individuals with average negative ED), and β_{11} represents the difference between the slope for negative ED on days with a negative event and the slope for negative ED on days without a negative event. Random slopes (r_{1i}) were tested for significance by applying a backward procedure, and only significant random slopes were retained in the model (Snijders & Bosker, 2012). All multilevel regression models were created with the R package lme4 (Bates, Mächler, Bolker, & Walker, 2015), and *p* values were computed with the R package lmerTest (Kuznetsova, Brockhoff, & Christensen, 2015). The within- and between-person correlations of Level 1 variables were computed in Mplus (Muthén & Muthén, 1998-2012).

Results and Discussion

Response Process (Specificity Index). On average, participants reported 2.55 (SD = 1.60) adjectives after a negative event. Momentary time pressure was unrelated to the number of adjectives (b = 0.04, t = 0.68, p = .497). Similar to a manipulation check in experimental studies, we additionally checked the plausibility of the valence of the adjectives that were reported in response to negative daily events: 93% of the adjectives were negative, 3% were neutral, and 3% were positive—thereby confirming that individuals complied with the task.[4]

Descriptive Statistics. Means, standard deviations, and correlations between the main variables are displayed in Table 1. Note that the mean for the specificity index of negative ED represents the proportion of specific words used. That is, the mean levels of negative ED were generally rather high, but nonetheless, individual differences were captured. There were neither sex differences in the specificity index of negative ED, $t(33.67) = -1.31$, $p = .198$, $M_{male} = 0.80$, $M_{female} = 0.89$, nor differences related to highest educational achievement, $F(3, 43.02) = 1.65$, $p = .191$. Age was also unrelated to the specificity index of negative ED ($r = .15$, $p = .111$).

Predictive Validity. The specificity index of negative ED was unrelated to self-reported ED ($r = .12$, $p = .224$). As can be seen in Table 1, the expected negative relations between ED and the alexithymia facets were found for self-reported ED (DIF: $r = -.31$, $p < .001$; DDF: $r = -.48$, $p < .001$). However, the specificity index of negative ED was unrelated to the alexithymia facets.

Next, we analyzed emotion regulation. The relations of the emotion regulation strategies with aggregated daily well-being (positive for reappraisal, negative for suppression, see Table 1) supported our expectation that reappraisal would be seen as rather functional and suppression as rather dysfunctional for well-being. As expected, higher self-reported ED was related to the use of more reappraisal ($r = .22, p = .022$) and less suppression ($r = -.19, p = .043$). The specificity index of negative ED was not related to reappraisal or suppression.

To analyze the association between ED and well-being, we first scrutinized the between-person correlations between the ED measures and daily well-being: As can be seen in Table 1, the specificity index of negative ED was related to higher daily life satisfaction ($r = .26, p = .014$) and more pleasant mood ($r = .26, p = .015$). This finding is in line with previous studies that found that negative ED was related to more positive outcomes (e.g., Erbas et al., 2014; Feldman Barrett et al., 2001). The global self-report measure of ED was unrelated to daily well-being indicators (see Table 1).

Subsequently, we conducted multilevel analyses for the daily well-being indicators as dependent variables, which allowed us to apply a more differentiated analysis of the ED measures' predictive validity. In particular, these analyses enabled us to test the hypothesis that negative ED would be more predictive of daily well-being on days with (vs. without) a negative event. As can be seen in Table 2 (Model 1), the specificity index of negative ED was associated with higher daily life satisfaction on days without a negative event (significant β_{01}), and the nonsignificant cross-level interaction term (β_{11}) indicated that the negative ED slope did not differ between the two types of days (i.e., with vs. without a negative event). As can be seen from the simple slopes in Figure 1 (Panel A1), higher negative ED was related to higher satisfaction on both types of days. When momentary pleasant–unpleasant mood was analyzed as the dependent variable (Model 2 in Table 2), a similar picture arose. Negative ED was related to higher pleasant–unpleasant mood on days without a negative event (significant β_{01}), and the negative ED slope did not differ between the two types of days (nonsignificant β_{11}; see also Figure 1, Panel B1). Taken together, negative ED seemed to be important for well-being, and this finding applied to both types of days.

It is interesting that the present findings highlight differences not only between different measurement approaches (direct vs. indirect measures of ED) but also between different approaches that are used to measure relevant outcomes (global valence-unspecific trait measures vs. valence-specific state measures gathered in daily life). The only significant relations with adaptive outcomes that the self-report measure of ED showed were (small to moderate) relations with global trait measures (reappraisal, suppression, and alexithymia facets). Self-reported ED was unrelated to (ecologically valid) quasi-trait measures that were formed by aggregating repeated state measures of well-being in daily life. The specificity index of negative ED, on the other hand, demonstrated relations with state measures but not with global trait measures of adaptive psychological functioning. This pattern of results can be interpreted as demonstrating that common method variance shared between global trait self-reports (e.g., beliefs about oneself) might be responsible for the associations found between self-reports. It would be interesting to scrutinize whether the specificity index of ED is more closely related to emotion-regulation measures when the latter are assessed in daily life instead of via global trait measures.

One limitation of Study 1 is that we did not control for verbal ability. Verbal ability might play an important role when participants describe their affective state in their own words and might be a confounding variable (Labouvie-Vief et al., 1989). Moreover, we measured emotion regulation only with global self-reports but not with (potentially) more ecologically valid measures of daily emotion regulation. Finally, the novel specificity index of ED was not compared with an established performance-based measure of ED.

Study 2

In Study 2, we sought to replicate the findings from Study 1 and extend Study 1 in important ways: First, we included a measure of verbal ability. Second, we assessed emotion regulation not only via global self-report but also with daily diary measures. Third, a subsample of participants completed a second daily diary phase during which closed-ended emotion scales were assessed. This enabled us to compute our new specificity index and the previously used ICC measure in the same subsample and analyze the relation between the two ED measures.

Method

Participants. Participants were recruited via e-mail, posts on German online forums, and flyers that were distributed in different towns. They participated voluntarily and had the chance to win different vouchers (10-50 €). Because this study involved a (voluntary) second daily diary phase to additionally assess closed-ended emotion reports (to be able to compute ICCs), we aimed for a higher sample size than in Study 1. We estimated that 50% of the participants might agree to continue with the second daily diary phase, and hence, aimed to recruit at least 200 participants. A total of 209 individuals from the general population filled out the initial online questionnaire and took part in the daily diary phase that followed. Data from 10 individuals were excluded from the analyses because they spoke a language other than German as their native tongue. Data from nine individuals were excluded because they reported no negative events

Table 2. Multilevel Models (Fixed Effects) Predicting Well-Being by Emotion Differentiation (Study 1).

Outcome predictor	Coefficient	Estimate (*SE*)	95% CI	*df*	*t*
Model 1: Daily life satisfaction					
Intercept	β_{00}	3.16 (0.03)	[3.09, 3.23]		
Neg. ED	β_{01}	0.49 (0.15)	[0.19, 0.78]	131.2	3.23**
Neg. event	β_{10}	−0.38 (0.03)	[−0.44, −0.33]	1752.6	−13.60***
Neg. ED × Neg. event	β_{11}	0.01 (0.14)	[−0.27, 0.29]	1754.9	0.05
Model 2: Momentary P-U mood					
Intercept	β_{00}	4.85 (0.07)	[4.71, 4.98]		
Neg. ED	β_{01}	0.85 (0.29)	[0.27, 1.42]	126.7	2.89**
Neg. event	β_{10}	−0.61 (0.05)	[−0.70, −0.52]	1732.9	−13.25***
Neg. ED × Neg. event	β_{11}	0.20 (0.23)	[−0.25, 0.65]	1735.6	0.87

Note. CI = confidence interval; *df* = degrees of freedom; Neg. ED = specificity index of negative emotion differentiation (grand mean centered); Neg. event = dummy variable that coded whether there was a negative event during the day (reference category: no negative event); P-U mood = momentary pleasant−unpleasant mood.
*$p < .05$. **$p < .01$. ***$p < .001$.

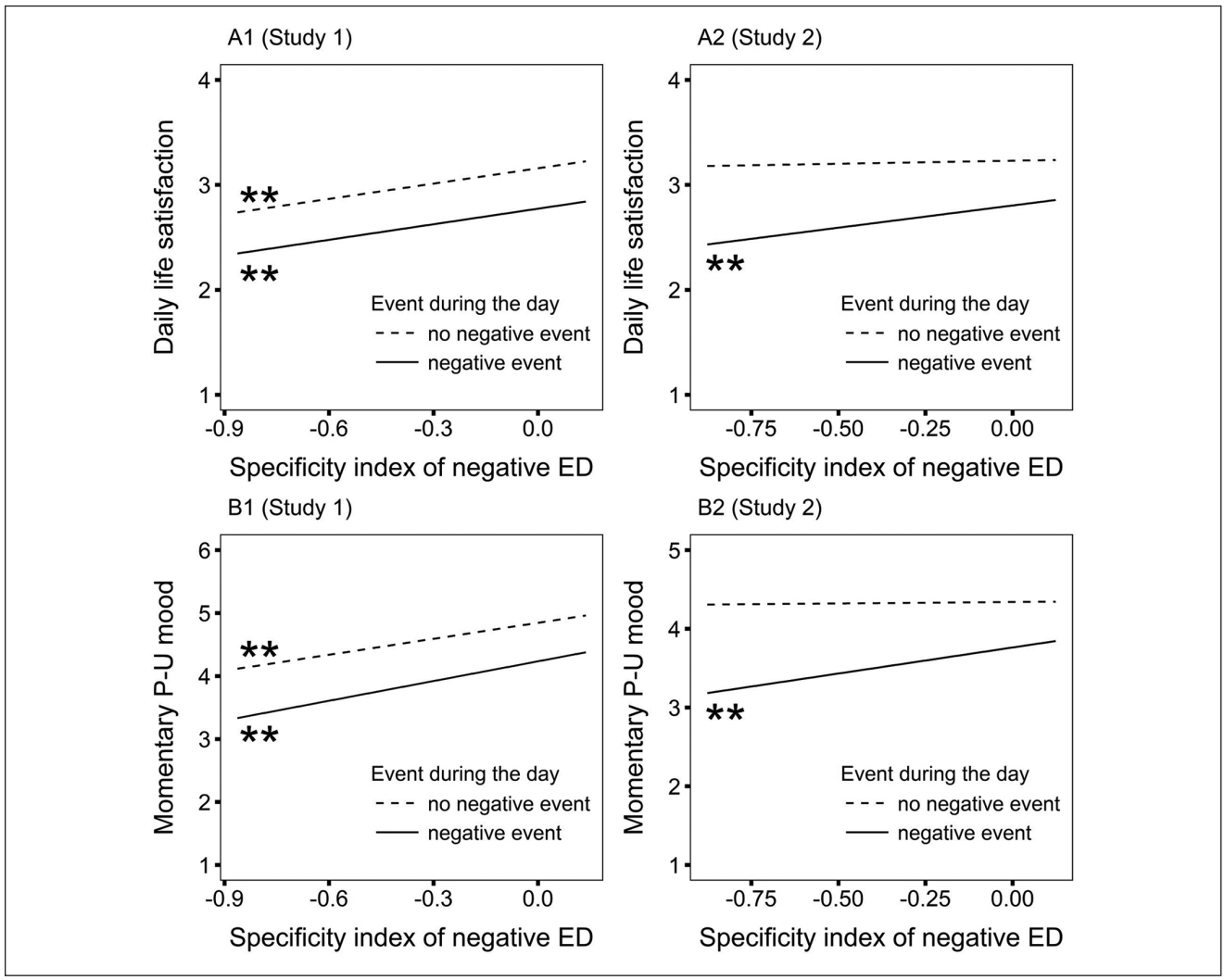


Figure 1. Simple slopes for the relation between negative emotion differentiation (specificity index) and daily well-being on different types of days.
Results from Study 1 in Column 1; results from Study 2 in Column 2. Statistically significant simple slopes are tagged with asterisks (*$p < .05$; **$p < .01$).

across the entire study. Thus, the final sample consisted of 190 individuals; 74% were female, and the mean age of the sample was 40.36 years ($SD = 16.01$). Participants' highest educational achievement was distributed as follows: 8% had no professional achievement or were in training, 27% had completed an apprenticeship, 19% had a polytechnic degree, and 46% had a university degree.

Procedure. The procedures used in Studies 1 and 2 were similar. One difference was that after the first daily diary phase in Study 2, participants were asked to complete another 3-week daily diary phase. The first phase was used to assess emotional experience with an open-ended format (to compute the specificity index of ED), and the second phase was used to assess emotional experience with a closed-ended format (to compute the ICC measure of ED). Compliance was good: Participants completed $M = 16.47$ ($SD = 4.79$) of the 21 questionnaires (one per day) during the first daily diary phase (range: 2-21). The number of negative events that were reported ranged from 1 to 21 ($M = 7.08$, $SD = 4.66$). One hundred and twelve persons decided to continue to participate in the second daily diary phase. Similar to the approach used by Starr, Hershenberg, Li, and Shaw (2017), an ICC was calculated for a participant if he or she had reported at least four negative events during the daily assessment phase. The data of 48 participants were excluded following this criterion.[5] Hence, the sample that was available for the analyses of the negative ICC measure consisted of 64 individuals who completed $M = 17.03$ ($SD = 3.55$) of the 21 questionnaires during the second daily diary phase (range: 4-21). The number of negative events that were reported ranged from 4 to 20 ($M = 9.14$, $SD = 4.34$).

Measures

Emotion differentiation–specificity index. The specificity index of negative ED was calculated in the exact same manner as in Study 1 with the data from the first daily diary phase. Reported emotion terms in Study 2 that had not been included in the dictionary ($N_{words} = 407$) were coded by two raters as specific or general. Interrater reliability for categorizing the new terms was good (Cohen's $\kappa = .83$). A last step before finalizing the dictionary was to add different spellings of the adjectives. The final dictionary consisted of 1,027 specific and 99 general adjectives.[6]

Emotion differentiation–ICC measure. The ICC measure of ED was calculated on the closed-ended emotional experience data from the second daily diary phase. To measure emotional experience during a negative event, six negative emotions (fear, anger, frustration, sadness, embarrassment, and boredom) were assessed. The items were rated on 5-point scales (0 = *not at all* to 4 = *very intense*). From these ratings, we computed performance-based indicators of negative ED. We computed the consistency ICC (ICC = 3; Erbas et al., 2014) for each participant across the emotion intensity variables.[7] Very low or even negative ICCs represent high ED.[8] For ease of interpretation, the ICCs were multiplied by −1 so that high values represented high ED.

Emotion differentiation–self-report. Again, we used the differentiation subscale from the RDEES (Kang & Shaver, 2004) as described in Study 1. Cronbach's α was .79.

Alexithymia. Alexithymia facets were again assessed with two subscales from the TAS-20 (Bagby et al., 1994). Cronbach's α was .74 for DIF and .75 for DDF.

Reappraisal and suppression (global self-report). Reappraisal and suppression were again assessed with the ERQ (Gross & John, 2003). Cronbach's α was .79 for reappraisal and .66 for suppression.

Daily reappraisal and suppression. During the first daily diary phase, suppression of negative emotions was measured with two items ("When I was feeling negative emotions, I was careful not to express them" from Kashdan & Steger, 2006; "When I felt negative emotions, people could easily see exactly what I was feeling" from Gross & John, 1995). Reappraisal was measured with two items ("I changed the way I thought about the situation I was in" from Kashdan & Steger, 2006; "I looked at things differently" from Totterdell & Parkinson, 1999). All items were answered on a 4-point response scale (1 = *strongly disagree* to 4 = *strongly agree*). For the two-item measures of emotion regulation, local (within-occasion) reliability (Buse & Pawlik, 1996) was estimated via the polychoric correlation between the items for each measurement occasion. The median polychoric correlation was .48 for suppression and .65 for reappraisal. Aggregate (between-occasions) reliability (Buse & Pawlik, 1996) was estimated by calculating the Pearson correlation between the two aggregated items (i.e., aggregated across days), which was .40 for suppression and .45 for reappraisal. These interitem correlations fell within a range that is recommended for correlations between questionnaire items that measure the same construct (e.g., Clark & Watson, 1995).

Daily life satisfaction. To measure daily life satisfaction at the end of each day, we used the same measure as in Study 1. The median polychoric correlation between the two items was .95. The Pearson correlation between the two aggregated items (i.e., aggregated across occasions) was .86.

Momentary pleasant–unpleasant mood. At each occasion, participants rated their momentary pleasant–unpleasant

Table 3. Means, Standard Deviations, Bivariate Correlations, and Partial Correlations for Main Variables (Study 2).

Variable	1	2	3	4	5	6	7	8	9	10
Emotion differentiation										
1. $ED_{self\text{-}report}$	—									
2. Negative $ED_{specificity}$	−.03 (−.03)	—								
Alexithymia										
3. DIF	−.18*	−.21** (−.20**)	—							
4. DDF	−.25***	−.06 (−.06)	.56***	—						
Emotion regulation										
5. Reap	.31***	.01 (.02)	−.26***	−.21**	—					
6. Aggr. daily Reap	.011	.19* (.18*)	−.06	−.014	.23**	—		.09**[a]	.08*[a]	.10***[a]
7. Supp	−.10	−.04 (−.04)	.29***	.51***	−.04	−.11	—			
8. Daily Supp	−.09	.04 (.00)	−.05	−.09	.10	.20	.17†	—	.15***[a]	.17***[a]
Daily well-being										
9. Daily LS	−.03	.04 (−.02)	−.24**	−.08	.25**	.014	−.19†	.31**	—	.58***[b]
10. P-U mood	−.00	.02 (.00)	−.37***	−.08	.28***	.16†	−.15†	.14	.71***	—
M	2.90	0.88	1.72	1.85	4.63	2.05	3.20	2.44	3.05	4.10
SD	0.45	0.19	0.45	0.52	1.01	0.51	1.10	0.49	0.26	0.44

Note. Correlation coefficients in parentheses are partial correlations controlling for verbal ability. Between-person correlations ($N_{Persons}$ = 190) are presented below the diagonal; within-person correlations among daily measures ([a]N_{Days} = 1,346; [b]N_{Days} = 3,130) are presented above the diagonal. $ED_{self\text{-}report}$ = Global self-report of emotion differentiation; Negative $ED_{specificity}$ = specificity index of negative emotion differentiation; DIF = Difficulty identifying feelings; DDF = difficulty describing feelings; Reap = reappraisal; Supp = suppression; Daily LS = daily life satisfaction; P-U mood = momentary pleasant–unpleasant mood.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

mood with three items from the short version of the Multidimensional Mood Questionnaire (Steyer, Schwenkmezger, Notz, & Eid, 1997). In contrast to Study 1, where a bipolar response format was used, the three items (unwell, happy, bad) were answered on the original monopolar 5-point intensity scale (1 = *not at all* to 5 = *very*). The items unwell and bad were reverse-coded so that higher values represented more pleasant mood. The within-person-level α (Geldhof et al., 2014) was .81, and the between-persons-level α was .85.

Momentary time pressure. Momentary time pressure was measured with the same single item as in Study 1. Mean time pressure was 1.77 (*SD* = 0.53).

Verbal ability. Verbal ability was assessed with the German Multiple Choice Vocabulary Test (Mehrfachwahl−Wortschatz−Intelligenztest [MWT-B]; Lehrl, 1999). The items consist of lines with five words each. One word is an authentic word from a German dictionary, whereas four are fictitious (meaningless letter combinations, which were similar to the target word). The task was to identify the authentic word in each line. The test consists of 37 items with increasing difficulty. In previous studies, the MWT-B showed high test−retest reliability and high parallel test reliability (for an overview, see Lehrl, Triebig, & Fischer, 1995) as well as high convergent correlations with other verbal intelligence measures (e.g., *r* = .55 to .59 with the HAWIE-R Verbal IQ, Satzger, Fessmann, & Engel, 2002). In our study, the mean sum score was 31.52 (*SD* = 3.79), and the reliability was .81 (estimated with the Kuder−Richardson Formula 20; Kuder & Richardson, 1937).

Results and Discussion

Response Process (Specificity Index). On average, participants reported 2.62 (*SD* = 1.40) adjectives after a negative event. Higher momentary time pressure was unrelated to the number of adjectives reported after a negative event ($b = -0.02$, $t = -0.39$, $p = .698$). Again, we checked the plausibility of the valence of the adjectives reported after negative daily events: After negative events, 91% of the adjectives were negative, 4% were neutral, and 5% were positive—thereby confirming that individuals complied with the task.[9] The means, standard deviations, and correlations between the main variables are displayed in Table 3.

Descriptive Statistics. Means, standard deviations, and correlations between the main variables are displayed in Table 3. There were neither sex differences in the specificity index of negative ED, $t(188) = 0.68$, $p = .499$, $M_{male} = 0.89$, $M_{female} = 0.87$, nor differences related to highest educational achievement, $F(3, 51.09) = 2.20$, $p = .100$. Age was also unrelated to the specificity index of negative ED ($r = -.01$, $p = .858$).

Table 4. Multilevel Models (Fixed Effects) Predicting Emotion Regulation and Well-Being by Emotion Differentiation (Study 2).

Outcome predictor	Coefficient	Estimate (*SE*)	95% CI	*df*	*t*
Daily emotion regulation					
Model 1: Suppression					
Intercept	β_{00}	2.41 (0.03)	[2.34, 2.47]		
Neg. ED	β_{01}	0.03 (0.21)	[−0.37, 0.44]	317.10	0.17
Model 2: Reappraisal					
Intercept	β_{00}	2.03 (0.04)	[1.96, 2.10]		
Neg. ED	β_{01}	0.40 (0.21)	[−0.02, 0.82]	302.02	1.86†
Daily well-being					
Model 3: Daily life satisfaction					
Intercept	β_{00}	3.23 (0.02)	[3.19, 3.27]		
Neg. ED	β_{01}	0.06 (0.10)	[−0.14, 0.26]	239.4	0.56
Neg. event	β_{10}	−0.43 (0.02)	[−0.47, −0.39]	3125.8	−20.68***
Neg. ED × Neg.	β_{11}	0.37 (0.13)	[0.12, 0.62]	3111.7	2.88**
Event					
Model 4: Momentary P-U mood					
Intercept	β_{00}	4.34 (0.03)	[4.28, 4.40]		
Neg. ED	β_{01}	0.04 (0.15)	[−0.25, 0.32]	169.5	0.25
Neg. event	β_{10}	−0.58 (0.03)	[−0.64, −0.51]	168.8	−17.52***
Neg. ED × Neg.	β_{11}	0.62 (0.20)	[0.24, 1.01]	244.0	3.18**
Event					

Note. CI = confidence interval; *df* = degree of freedom; Neg. ED = specificity index of negative emotion differentiation (grand mean centered); Neg. event = dummy variable that coded whether there was a negative event during the day (reference category: no negative event); P-U mood = momentary pleasant–unpleasant mood.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Analysis of Dropout. After the first daily diary phase, participants decided whether they wanted to continue to participate. To check whether there was systematic dropout after the first daily diary phase, we compared participants who dropped out versus participants who decided to continue on the main study variables. Independent-samples *t* tests revealed no differences in negative ED (specificity index), self-reported ED, (aggregated) pleasant–unpleasant mood, (aggregated) daily life satisfaction, or verbal ability ($|t| \leq 1.06, p > .290$). Thus, we concluded that there was no systematic dropout after the first daily diary phase.

Relations Between ED Measures. Self-reported ED was unrelated to the specificity index of negative ED (see Table 3, $r = -.03$, $n = 190$; $p = .666$, 95% confidence interval [CI] = [−.17, .11]) and to the ICC measure of negative ED ($r = -.23$, $n = 64$; $p = .071$, 95% CI = [−.45, .02]). The specificity index of negative ED was unrelated to the ICC measure ($r = -.03, n = 64, p = .822$, 95% CI = [−.27, .22]). Note that the correlations with the ICC measure were based on only a subsample.[10] As computing ICCs requires a minimum number of reported events, the sample size was strongly reduced. Hence, these correlations should be interpreted cautiously.

Associations Between ED Measures and Verbal Ability. Verbal ability was unrelated to the specificity index of negative ED ($r = .10$, $p = .162$, 95% CI = [−.04, .24]).[11] Moreover, verbal ability was unrelated to the other ED measures ($r_{\text{self-report}} = -.05$, $p = .494$, 95% CI = [−.19, .09]; $r_{\text{ICC}} = -.09$, $p = .482$, 95% CI = [−.33, .16]).

Predictive Validity. The expected negative relations between ED and the alexithymia facets were found for self-reported ED (DIF: $r = -.18$, $p = .014$; DDF: $r = -.25$, $p < .001$; Table 3), in line with the Study 1 results and Kang and Shaver's (2004) findings. The specificity index of negative ED was related to higher DIF ($r = -.21$, $p = .003$) but unrelated to DDF. This replicates Erbas et al.'s (2014, Study 2) findings.

With respect to the relations between ED and emotion regulation, the results were mixed (see Table 3): Self-reported ED showed the expected relations with emotion regulation only for global self-reported reappraisal but not for other regulation measures (global self-reported suppression, daily suppression, and reappraisal after negative events). The specificity index of negative ED showed a small positive correlation with daily reappraisal but was unrelated to the other emotion-regulation measures. Multilevel models with negative ED predicting emotion regulation in daily life confirmed these results (see Table 4, upper part). The specificity index of negative ED was positively associated with daily reappraisal ($\beta_{01} = 0.40$, $t = 1.86$, one-tailed $p = .032$).[12]

As in Study 1, the relations between global emotion regulation and well-being supported our expectation that reappraisal could be seen as rather functional and suppression as rather dysfunctional for well-being (see Table 3). However, daily suppression of emotions after a negative event was related to higher daily life satisfaction.

To analyze the relation between ED and well-being, we first inspected the between-person correlations between the ED measures and daily well-being. The correlations were very small and nonsignificant. Multilevel analyses allowed us to analyze this relation without aggregating the daily measures by hand and to test whether ED showed differential relations with well-being, depending on whether a negative event was experienced that day. As can be seen in Table 4 (Models 3 and 4), the specificity index of negative ED was unrelated to daily life satisfaction and momentary pleasant–unpleasant mood on days without a negative event (nonsignificant β_{01}), and the significant cross-level interaction terms (β_{11}) indicated that the negative ED slope was higher on days with a negative event. As can be seen from the simple slopes in Figure 1 (Panels A2 and B2), higher negative ED was related to higher satisfaction and more pleasant–unpleasant mood on days with a negative event.[13] These results supported the expectation that negative ED is most beneficial on days that present a challenge to individuals' well-being compared with less challenging days.

General Discussion

The goal of the present research was to develop a novel behavioral index of assessing ED that closely matched the definition of the construct of ED (see, e.g., Feldman Barrett, 2006) due to its focus on the use of self-selected emotion words in daily life. We used a multimethod assessment strategy by combining two different performance-based measures and self-reported ED. Relations between these ED measures were analyzed. Moreover, the predictive validity of ED measures with respect to emotion regulation and subjective well-being was investigated in two daily diary studies, and we tested whether verbal ability could account for variance in the specificity index of ED.

Comparative Evaluation of Different Approaches to Measuring ED

We used three different approaches to measure ED: a self-report, the novel specificity index that relied on the proportion of specific emotion terms used by participants when they described their emotional experiences in their own words, and (in a subsample) an established behavioral measure that was based on the covariation of different emotions over time (ICC measure).

An advantage of self-report measures of ED is that they are easy to administer and can easily be included in different kinds of studies (e.g., experimental, cross-sectional, or longitudinal designs). The disadvantages of direct (self-report) measures are that they rely on participants' willingness and ability to report on the construct of interest. They can be plagued by memory biases (Conner & Feldman Barrett, 2012) or social desirability bias (e.g., Kluemper, 2008). A specific disadvantage of the (only existing) self-report measure of ED by Kang and Shaver (2004) is that it does not distinguish between positive and negative ED. Indirect (behavioral) measures of ED are an alternative to self-report measures. Although performance-based indices of ED (e.g., the ICC measure) use (repeated) self-reports of emotional experience, they represent an indirect (behavioral) measure in the sense that individuals do not rate their standing on the dimension of ED; instead, ED is inferred from the intraindividual covariation of different emotion terms over time. In a similar way, our novel specificity index of ED requires self-reports of emotional experience, but it is a behavioral (i.e., indirect) method in the sense that the emotion words are categorized by raters. Participants do not know how their data will be processed. Thus, compared with self-reports, it is less likely that participants will intentionally bias their ED values in a certain direction. Compared with self-reports, which require only a single measurement occasion, previously suggested performance-based indices of ED require repeated measurements of affect (in the context of experience sampling or laboratory studies). In principle, the novel specificity index of ED could also be applied to a single measurement occasion, but in order to tap dispositional individual differences in ED and to increase reliability, we used multiple measurement occasions per person to calculate the specificity index.

When comparing previous performance-based indices (e.g., the ICC measure) with the novel specificity index (proportion of specific affective adjectives), the most important advantage of the novel method is that it better matches the definition of high ED as the ability to use specific emotion terms to describe an affective state (e.g., Feldman Barrett, 2006). Moreover, the specificity method does not equate the co-occurrence of emotions of the same valence with a lack of granularity. Although we did not use the verbal descriptions of experiences that we collected to code the degree to which individuals experienced mixed emotions of the same valence (e.g., anger and sadness) during an event or the degree to which an event evoked mixed emotions of opposite valence (e.g., joy and sadness), this could in principle be done by additionally coding adjectives from the specific category with respect to their emotion-specific content (cf. Grossmann & Ellsworth, 2017; Grossmann et al., 2016). We believe that by asking participants to report affect terms that come to their minds, the

new specificity method of ED circumvents problems inherent in previous performance-based ED indices. With the ICC measure, for instance, it is not possible to detect *which* terms participants would have used because a closed-ended list of emotion terms is provided by the researchers (Kashdan et al., 2015). A drawback is that our new specificity index of ED necessitates the coding of reported terms as either specific or global. As demonstrated in the present studies, however, rater agreement on the classification of terms was good, thus attesting to the scoring procedure's objectivity (reliability).

In Study 2, we used both the specificity index and the ICC measure and found that neither the specificity index nor the ICC measure was related to self-reported ED. Only one publication described the relation between self-reported ED and a performance-based measure of ED (unpublished raw data from Feldman Barrett, 2001, cited in Lindquist & Feldman Barrett, 2008). They reported that "the RDEES also failed to correlate with indices of emotional granularity computed over a 1-month period" (Lindquist & Feldman Barrett, 2008, p. 523). Research involving indirect measures of attitudes and personality traits has shown that convergent correlations between indirect (behavioral) measures and direct (self-report) measures of the same target construct are typically low to moderate (e.g., Robinson & Neighbors, 2006). It should again be noted that the sample size for calculating the ICC measure was greatly reduced because we set a cutoff for the minimum number of reported events. Future studies should ensure that participants report more daily events or that the study starts with a much larger sample of participants so that the final sample for calculating the ICC measure is sufficiently large.

Relations Between Emotion Differentiation Measures and Verbal Ability

We found that all the ED indices that we used (specificity index, ICC, self-report) were unrelated to verbal ability, thereby demonstrating discriminant validity. In previous research, the LEAS as a measure of emotional awareness showed low to moderate correlations with verbal ability measures (e.g., Ciarrochi, Caputi, & Mayer, 2003; Feldman Barrett, Lane, Sechrest, & Schwartz, 2000; Lane et al., 1998), but relations with criterion variables remained stable when verbal ability was controlled for (e.g., Feldman Barrett et al., 2000).

Predictive Validity of Emotion Differentiation Measures

In both studies, self-reported ED was negatively related to the alexithymia facets and positively related to the global self-reports of reappraisal (and in Study 1 negatively related to the global self-reports of suppression). However, self-reported ED was unrelated to the daily emotion-regulation and well-being measures. This may be interpreted as showing that global self-report measures of emotion-related traits mainly tap individuals' beliefs about themselves (e.g., Robinson & Clore, 2002). Another explanation may be that the daily measures of well-being and emotion regulation that we used were flawed. However, the daily measures demonstrated adequate reliability and convergent relations with global self-reports of the use of these regulation strategies.

For our novel specificity index of ED, the predictive relations with well-being indicators were consistent with theoretical accounts of negative ED (e.g., Kashdan et al., 2015). In particular, in both studies, negative ED demonstrated positive relations with cognitive and affective well-being in daily life—in Study 1, this relation held across different types of days, and in Study 2, this relation held on days when a negative event occurred. Hence, the hypothesis that the adaptive value of (negative) ED is stronger in situations when the need for negative emotion regulation is high (e.g., Feldman Barrett et al., 2001; Kashdan et al., 2010; Pond et al., 2012) received partial support. This result is also in line with findings from Starr et al. (2017, Study 2) who found that higher negative ED was related to lower depressed mood on days with more frequent negative experiences.

The relations of ED with alexithymia facets and the use of reappraisal and suppression were nonsignificant in both studies or inconsistent between the two studies. One explanation for this might be that the adaptive value of negative ED may be especially evident in (sub)clinical populations (and not in the healthy population that we studied). Many studies that have provided empirical evidence for the functionality of ED have been based on clinical samples (e.g., Demiralp et al., 2012; Erbas et al., 2013; Selby et al., 2014; Tomko et al., 2015). Hence, the lack of relation between ED and emotion regulation might be due to our samples, which were not particularly stressed, as indicated by low levels of time pressure and high overall levels of well-being. In other samples of "healthy" adults, a few studies have found correlational patterns that indicate the functionality of ED for emotion regulation and coping (e.g., Feldman Barrett et al., 2001; Tugade et al., 2004), but findings of no relations with well-being measures have also been reported (e.g., Grühn et al., 2013). Future studies might further investigate the validity of the specificity index of ED either in "healthy" samples during high stress periods (e.g., during examination periods) or in (sub)clinical samples. Another reason that the results on the link between ED and emotion regulation were mixed in our studies (and in previous research) could be that the adaptiveness of a specific regulation strategy varies by context. In particular, the benefits of specific regulation strategies such as reappraisal might depend on the extent

to which the strategy is used in a context-sensitive manner (Aldao, Sheppes, & Gross, 2015). A challenge for future research might be to test whether ED facilitates the context-sensitive selection of regulation strategies.

Limitations

Because of our use of a correlational approach, the present findings cannot provide information about causal relationships. The positive relations of negative ED with life satisfaction and pleasant–unpleasant mood, for instance, could be interpreted in two ways: Higher negative ED might facilitate affect regulation and consequently lead to higher daily well-being. Alternatively, individuals with high daily well-being might have more resources and more motivation to reflect on their emotional state, and this might result in higher negative ED.

The present studies consisted of 3 weeks (Study 1) or 2 × 3 weeks (Study 2) with assessments taken only one time per day. It is possible that there were times when participants did not experience an emotion-eliciting event during the day. As we did not want to induce participants to report on events that did not happen, the reporting of an event was voluntary. Despite high compliance rates, we could not ensure that each participant would report a sufficient number of events, which were the basis for the calculation of ED scores. Especially in the second part of the second study, more occasions per person would have been valuable, which would have resulted in a larger sample size for the analyses on the ICC measure of ED. Future studies should systematically investigate how many data points (i.e., reported events) are needed to obtain a reliable person index of ED.

Conclusion

The present research is the first to follow recent recommendations (Kashdan et al., 2015) to use an open-response format to assess individual differences in ED. Although further validation of the novel specificity index is essential, we hope that the present findings will encourage researchers to integrate the proportion of specific affective terms as an ED index in their studies.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Participants also completed similar questions with respect to a positive daily event. Because of page restrictions, we refrained from including results on positive ED in the present article.
2. The dictionary can be downloaded from https://osf.io/6mk5u/?view_only=e1a828ebd77b4b74b1c0c841a64f1787 (doi: 10.17605/OSF.IO/6MK5U)
3. Data, R code, and Mplus syntax can be downloaded from https://osf.io/6mk5u/?view_only=e1a828ebd77b4b74b1c0c841a64f1787 (doi: 10.17605/OSF.IO/6MK5U)
4. When including only those adjectives that matched the valence of the event (i.e., negative and neutral words after negative events) in the ED index, we found a nearly perfect correlation between the reported ED index and the "valence-matching" ED index ($r = .99$, $p < .001$).
5. When at least 10 measurement occasions were used as the cutoff criteria, the results were very similar.
6. The dictionary can be downloaded from https://osf.io/6mk5u/?view_only=e1a828ebd77b4b74b1c0c841a64f1787 (doi: 10.17605/OSF.IO/6MK5U)
7. Shrout and Fleiss (1979) proposed three different ICCs that they named ICC 1 (absolute agreement), ICC 2, and ICC 3 (consistency). Each ICC makes certain assumptions (Shrout & Fleiss, 1979). Most studies have chosen the absolute agreement ICC (e.g., Boden et al., 2013; Grühn et al., 2013; Kashdan & Farmer, 2014; Pond et al., 2012; Tugade et al., 2004). ICC 1 assumes that for each measurement occasion, different emotions are rated. In contrast, ICCs 2 and 3 assume that the same emotions are rated at each measurement occasion. ICC 2 would be the right choice if the emotions were randomly selected from a population of emotions. Emotion terms cannot be treated as interchangeable but should be selected as the result of theoretical considerations, so ICC 3 should be preferred. Erbas et al. (2014) were the first authors to use this consistency ICC. Another difference between ICC 1 and ICC 3 is that the mean intensity of different emotions influences ICC 1 but not ICC 3. A high ICC 1 can be achieved only if the means of the different emotions are similar. As the similarity of mean intensities of different emotions should not influence the measure of ED, this is another reason to choose ICC 3.
8. For some of the participants, the ICC measure was negative. Analyses with negative ICCs set to 0 (which is suggested as an alternative practice in the field of interrater reliability; Baldwin, Murray, & Shadish, 2005) produced similar results.
9. When including only those adjectives that matched the valence of the event (negative and neutral words after negative events) in the ED index, we found a nearly perfect correlation between the reported ED index and the "valence-matching" ED index ($r = .98$, $p < .001$).
10. The correlations with the ICC measure were similar when participants with at least 10 reported events were included (self-reported ED: $r = -.02$, $n = 24$, $p = .908$; specificity index of negative ED: $r = -.08$, $n = 24$, $p = .694$).
11. There might be individuals who use a very small set of specific adjectives over and over again. To find out whether high specificity scores might result from using a small number of

unique adjectives over time, we additionally calculated the number of unique specific adjectives per person (divided by the number of reported events) and correlated this score with verbal ability and with our specificity index of ED. A larger number of unique specific adjectives was related to higher verbal ability ($r = .17$, $p = .020$). Moreover, it was positively related to the specificity index of ED (Study 1: $r = .39$, $p < .001$; Study 2: $r = .40$, $p < .001$). That is, individuals with high values on the specificity indices reported a larger number of different (unique) specific emotion terms. We controlled only for the verbal ability measure in Study 2 because it was unclear whether controlling for the number of unique specific adjectives would partial out a crucial part of individual differences in ED.

12. The results on ED and emotion regulation remained the same when we controlled for verbal ability.
13. The results on ED and daily life satisfaction or pleasant–unpleasant mood remained the same when we controlled for verbal ability.

ORCID iD

Charlotte Ottenstein https://orcid.org/0000-0003-0592-765X

References

Aldao, A., Sheppes, G., & Gross, J. J. (2015). Emotion regulation flexibility. *Cognitive Therapy and Research*, *39*, 263-278. doi:10.1007/s10608-014-9662-4

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*, 23-32. doi:10.1016/0022-3999(94)90005-1

Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or type I errors? Problems with the analysis of data from group-administered treatments. *Journal of Consulting and Clinical Psychology*, *73*, 924-935. doi:10.1037/0022-006X.73.5.924

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*. doi:10.18637/jss.v067.i01

Boden, M. T., & Thompson, R. J. (2015). Facets of emotional awareness and associations with emotion regulation and depression. *Emotion*, *15*, 399-410. doi:10.1037/emo0000057

Boden, M. T., Thompson, R. J., Dizén, M., Berenbaum, H., & Baker, J. P. (2013). Are emotional clarity and emotion differentiation related? *Cognition and Emotion*, *27*, 961-978. doi:10.1080/02699931.2012.751899

Boyd, R. (2014). *RIOT Scan*. Retrieved from http://riot.ryanb.cc/

Brans, K., Koval, P., Verduyn, P., Lim, Y. L., & Kuppens, P. (2013). The regulation of negative and positive affect in daily life. *Emotion*, *13*, 926-939. doi:10.1037/a0032400

Buse, L., & Pawlik, K. (1996). Ambulatory behavioural assessment and in-field performance testing. In J. Fahrenberg & M. Myrtek (Eds.), *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 29-50). Seattle, WA: Hogrefe & Huber.

Ciarrochi, J., Caputi, P., & Mayer, J. D. (2003). The distinctiveness and utility of a measure of trait emotional awareness. *Personality and Individual Differences*, *34*, 1477-1490. doi:10.1016/s0191-8869(02)00129-0

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309-319. doi:10.1037/1040-3590.7.3.309

Conner, T. S., & Feldman Barrett, L. (2012). Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. *Psychosomatic Medicine*, *74*, 327-337. doi:10.1097/PSY.0b013e3182546f18

Demiralp, E., Thompson, R. J., Mata, J., Jaeggi, S. M., Buschkuehl, M., Feldman Barrett, L., . . . Jonides, J. (2012). Feeling blue or turquoise? Emotional differentiation in major depressive disorder. *Psychological Science*, *23*, 1410-1416. doi:10.1177/0956797612444903

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

Erbas, Y., Ceulemans, E., Boonen, J., Noens, I., & Kuppens, P. (2013). Emotion differentiation in autism spectrum disorder. *Research in Autism Spectrum Disorders*, *7*, 1221-1227. doi:10.1016/j.rasd.2013.07.007

Erbas, Y., Ceulemans, E., Koval, P., & Kuppens, P. (2015). The role of valence focus and appraisal overlap in emotion differentiation. *Emotion*, *15*, 373-382. doi:10.1037/emo0000039

Erbas, Y., Ceulemans, E., Pe, M. L., Koval, P., & Kuppens, P. (2014). Negative emotion differentiation: Its personality and well-being correlates and a comparison of different assessment methods. *Cognition and Emotion*, *28*, 1196-1213. doi:10.1080/02699931.2013.875890

Feldman Barrett, L. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition and Emotion*, *12*, 579-599. doi:10.1080/026999398379574

Feldman Barrett, L. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*, 20-46. doi:10.1207/s15327957pspr1001_2

Feldman Barrett, L., Gross, J. J., Christensen, T., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*, *15*, 713-724. doi:10.1080/02699930143000239

Feldman Barrett, L., Lane, R., Sechrest, L., & Schwartz, G. (2000). Sex differences in emotional awareness. *Personality and Social Psychology Bulletin*, *26*, 1027-1035. doi:10.1177/01461672002611001

Geisler, F. C. M., & Schröder-Abé, M. (2015). Is emotion suppression beneficial or harmful? It depends on self-regulatory strength. *Motivation and Emotion*, *39*, 553-562. doi:10.1007/s11031-014-9467-5

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*, 72-91. doi:10.1037/a0032138

Germain, C. L., & Kangas, M. (2015). Trait anger symptoms and emotion regulation: The effectiveness of reappraisal, acceptance and suppression strategies in regulating anger. *Behaviour Change*, *32*, 35-45. doi:10.1017/bec.2014.28

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expres-

sion, and physiology. *Journal of Personality and Social Psychology*, *74*, 224-237. doi:10.1037/0022-3514.74.1.224

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, *26*, 1-26. doi:10.1080/1047840X.2014.940781

Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences*, *19*, 555-568. doi:10.1016/0191-8869(95)00055-B

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. doi:10.1037/0022-3514.85.2.348

Grossmann, I., & Ellsworth, P. C. (2017). What are mixed emotions and what conditions foster them? Life-span experiences, culture and social awareness. *Current Opinion in Behavioral Sciences*, *15*, 1-5. doi:10.1016/j.cobeha.2017.05.001

Grossmann, I., Huynh, A. C., & Ellsworth, P. C. (2016). Emotional complexity: Clarifying Definitions and cultural correlates. *Journal of Personality and Social Psychology*, *111*, 895. doi:10.1037/pspp0000084.supp

Grühn, D., Lumley, M. A., Diehl, M., & Labouvie-Vief, G. (2013). Time-based indicators of emotional complexity: Interrelations and correlates. *Emotion*, *13*, 226-237. doi:10.1037/a0030363

Hay, E. L., & Diehl, M. (2011). Emotion complexity and emotion regulation across adulthood. *European Journal of Ageing*, *8*, 157-168. doi:10.1007/s10433-011-0191-7

Kang, S.-M., & Shaver, P. R. (2004). Individual differences in emotional complexity: Their psychological implications. *Journal of Personality*, *72*, 687-726. doi:10.1111/j.0022-3506.2004.00277.x

Kashdan, T. B., & Farmer, A. S. (2014). Differentiating emotions across contexts: Comparing adults with and without social anxiety disorder using random, social interaction, and daily experience sampling. *Emotion*, *14*, 629-638. doi:10.1037/a0035796

Kashdan, T. B., Feldman Barrett, L., & McKnight, P. E. (2015). Unpacking emotion differentiation: Transforming unpleasant experience by perceiving distinctions in negativity. *Current Directions in Psychological Science*, *24*, 10-16. doi:10.1177/0963721414550708

Kashdan, T. B., Ferssizidis, P., Collins, R. L., & Muraven, M. (2010). Emotion differentiation as resilience against excessive alcohol use: An ecological momentary assessment in underage social drinkers. *Psychological Science*, *21*, 1341-1347. doi:10.1177/0956797610379863

Kashdan, T. B., & Steger, M. F. (2006). Expanding the topography of social anxiety: An experience-sampling assessment of positive emotions, positive events, and emotion suppression. *Psychological Science*, *17*, 120-128. doi:10.1111/j.1467-9280.2006.01674.x

Kennedy-Moore, E., & Watson, J. (1999). *Expressing emotion: Myths, realities, and therapeutic strategies. Emotions and social behavior*. New York, NY: Guilford Press.

Kimhy, D., Vakhrusheva, J., Khan, S., Chang, R. W., Hansen, M. C., Ballon, J. S., . . . Gross, J. J. (2014). Emotional granularity and social functioning in individuals with schizophrenia: An experience sampling study. *Journal of Psychiatric Research*, *53*, 141-148. doi:10.1016/j.jpsychires.2014.01.020

Kirby, L. D., Tugade, M. M., Morrow, J., Ahrens, A. H., & Smith, C. (2014). Vive la différence: The ability to differentiate positive emotional experience and well-being. In M. M. Tugade, M. N. Shiota & L. D. Kirby (Eds.), *Handbook of positive emotions* (pp. 241-255). New York, NY: Guilford Press.

Kluemper, D. H. (2008). Trait emotional intelligence: The impact of core-self evaluations and social desirability. *Personality and Individual Differences*, *44*, 1402-1412. doi:10.1016/j.paid.2007.12.008

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*, 151-160.

Kuznetsova, A., Brockhoff, P. B., & Christensen, H. (2015). *lmerTest: Tests in linear mixed effects models. R package version 2.0-32*. Retrieved from http://CRAN.R-project.org/package=lmerTest

Labouvie-Vief, G., DeVoe, M., & Bulka, D. (1989). Speaking about feelings: Conceptions of emotion across the life span. *Psychology and Aging*, *4*, 425-437.

Lane, R., Quinlan, D. M., Schwartz, G., Walker, P., & Zeitlin, S. B. (1990). The levels of emotional awareness scale: A cognitive-developmental measure of emotion. *Journal of Personality Assessment*, *55*, 124-134. doi:10.1080/00223891.1990.9674052

Lane, R., Reiman, E. M., Axelrod, B., Yun, L.-S., Holmes, A., & Schwartz, G. (1998). Neural correlates of levels of emotional awareness: Evidence of an interaction between emotion and attention in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *10*, 525-535. doi:10.1162/089892998562924

Lehrl, S. (1999). *Manual zum MWT-B* [Manual for the MWT-B]. Balingen, Germany: Spitta.

Lehrl, S., Triebig, G., & Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*, *91*, 335-345. doi:10.1111/j.1600-0404.1995.tb07018.x

Lindquist, K. A., & Feldman Barrett, L. (2008). Emotional complexity. In M. Lewis, J. M. Haviland-Jones & L. Feldman Barrett (Eds.), *Handbook of emotions* (pp. 513-530). New York, NY: Guilford Press.

Lindquist, K. A., Gendron, M., Oosterwijk, S., & Feldman Barrett, L. (2013). Do people essentialize emotions? Individual differences in emotion essentialism and emotional experience. *Emotion*, *13*, 629-644. doi:10.1037/a0032283

Lischetzke, T., & Eid, M. (2017). The functionality of emotional clarity: A process-oriented approach to understanding the relation between emotional clarity and well-being. In M. D. Robinson & M. Eid (Eds.), *The happy mind: Cognitive contributions to well-being* (pp. 371-388). Dordrecht, Netherlands: Springer. doi:10.1007/978-3-319-58763-9_20

Lischetzke, T., Pfeifer, H., Crayen, C., & Eid, M. (2012). Motivation to regulate mood as a mediator between state extraversion and pleasant-unpleasant mood. *Journal of Research in Personality*, *46*, 414-422. doi:10.1016/j.jrp.2012.04.002

Luminet, O., Rimé, B., Bagby, R. M., & Taylor, G. (2004). A multimodal investigation of emotional responding in alexithymia. *Cognition and Emotion*, *18*, 741-766. doi:10.1080/02699930341000275

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

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. doi:10.1037/0022-3514.93.4.685

O'Toole, M. S., Jensen, M. B., Fentz, H. N., Zachariae, R., & Hougaard, E. (2014). Emotion differentiation and emotion regulation in high and low socially anxious individuals: An experience-sampling study. *Cognitive Therapy and Research*, *38*, 428-438. doi:10.1007/s10608-014-9611-2

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC2001*. Mahwah, NJ: Lawrence Erlbaum.

Plonsker, R., Gavish Biran, D., Zvielli, A., & Bernstein, A. (2016). Cognitive fusion and emotion differentiation: Does getting entangled with our thoughts dysregulate the generation, experience and regulation of emotion? *Cognition and Emotion*, *31*, 1286-1293. doi:10.1080/02699931.2016.1211993

Pond, R. S., Kashdan, T. B., DeWall, C. N., Savostyanova, A., Lambert, N. M., & Fincham, F. D. (2012). Emotion differentiation moderates aggressive tendencies in angry people: A daily diary analysis. *Emotion*, *12*, 326-337. doi:10.1037/a0025762

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*, 934-960. doi:10.1037//0033-2909.128.6.934

Robinson, M. D., & Neighbors, C. (2006). Catching the mind in action: Implicit methods in personality research and assessment. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 115-125). Washington, DC: American Psychological Association. doi:10.1037/11383-009

Satzger, W., Fessmann, H., & Engel, R. (2002). Liefern HAWIE-R, WST und MWT-B vergleichbare IQ-Werte? [The equivalence of three German vocabulary tests and the German version of the Wechsler Adult Intelligence Scale-Revised (HAWIE-R)]. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*, *23*, 159-170. doi:10.1024//0170-1789.23.2.159

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 527-561). New York, NY: Guilford Press.

Schwarz, N. (2012). Feelings-as-information theory. In P. A. M. Van Lange, A. W. Kruglanski & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 289-308). Los Angeles, CA: Sage. doi:10.4135/9781446249215.n15

Selby, E. A., Wonderlich, S. A., Crosby, R. D., Engel, S., Panza, E., Mitchell, J. E., . . . Le Grange, D. (2014). Nothing tastes as good as thin feels: Low positive emotion differentiation and weight-loss activities in anorexia nervosa. *Clinical Psychological Science*, *2*, 514-531. doi:10.1177/2167702613512794

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420-428. doi:10.1037/0033-2909.86.2.420

Sifneos, P. E. (1973). The prevalence of "alexithymic" characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics*, *22*, 255-262. doi:10.1159/000286529

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, CA: Sage.

Starr, L. R., Hershenberg, R., Li, Y. I., & Shaw, Z. A. (2017). When feelings lack precision: Low positive and negative emotion differentiation and depressive symptoms in daily life. *Clinical Psychological Science*, *5*, 613-631. doi:10.1177/2167702617694657

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1994). Testtheoretische Analysen des Mehrdimensionalen Befindlichkeitsfragebogen (MDBF) [Testtheoretical analysis of a multidimensional mood questionnaire (MDBF)]. *Diagnostica*, *40*, 320-328.

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF): Handanweisung* [*The multidimensional affect rating scale (MDBF): manual*]. Göttingen, Germany: Hogrefe.

Tomko, R. L., Lane, S., Pronove, L. M., Treloar, H. R., Brown, W. C., Solhan, M. B., . . . Trull, T. J. (2015). Undifferentiated negative affect and impulsivity in borderline personality and depressive disorders: A momentary perspective. *Journal of Abnormal Psychology*, *124*, 740-753. doi:10.1037/abn0000064

Totterdell, P., & Parkinson, B. (1999). Use and effectiveness of self-regulation strategies for improving mood in a group of trainee teachers. *Journal of Occupational Health Psychology*, *4*, 219-232. doi:10.1037/1076-8998.4.3.219

Trull, T. J., Lane, S., Koval, P., & Ebner-Priemer, U. W. (2015). Affective dynamics in psychopathology. *Emotion Review*, *7*, 355-361. doi:10.1177/1754073915590617

Tugade, M. M., Fredrickson, B. L., & Feldman Barrett, L. (2004). Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of Personality*, *72*, 1161-1190. doi:10.1111/j.1467-6494.2004.00294.x

Wolf, M., Horn, A. B., Mehl, M. R., Haug, S., Pennebaker, J. W., & Kordy, H. (2008). Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count [Computer-aided quantitative text analysis: Equivalence and reliability of the German adaptation of the Linguistic Inquiry and Word Count]. *Diagnostica*, *54*, 85-98. doi:10.1026/0012-1924.54.2.85

Zaki, L. F., Coifman, K. G., Rafaeli, E., Berenson, K. R., & Downey, G. (2013). Emotion differentiation as a protective factor against nonsuicidal self-injury in borderline personality disorder. *Behavior Therapy*, *44*, 529-540. doi:10.1016/j.beth.2013.04.008