A signal-detection approach to individual differences in negative feeling

Anat Karmon-Presser*, Nachshon Meiran
Department of Psychology and Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Israel
* Corresponding author.
E-mail address: anatkarm@post.bgu.ac.il (A. Karmon-Presser).

Abstract

Feeling is an important aspect of core personality traits and affective-style. Here we implemented a new signal-detection-theory based model for feeling generation, involving two parameters: report-criterion \((c)\), the level above which enough emotional evidence has gathered for intense feeling to appear, and evidence-differentiation \((d_a)\), the ability to emotionally differentiate between (negative) triggers of varying intensity. Results indicate that a low \(c\) was related to Neuroticism but not to affective-style, yet a low \(d_a\) was related to limited access to emotion regulation strategies, but not to personality traits.

Keywords: Psychology, Neuroscience

1. Introduction

People differ in how they perceive, experience and react emotionally (e.g., Barr et al., 2008; Suls and Martin, 2005; Bonanno and Burton, 2013). Here we focus on a prominent aspect in emotion, the subjective conscious experiences of the emotion, or feelings. Building on prior work (e.g., Barrett et al., 2007; Barbosa et al., 2016; Neufeld, 1975; Nielsen and Kaszniak, 2006), we (Karmon-Presser et al., 2018) have recently conceptualized feeling generation in terms of Signal-Detection Theory (SDT, Macmillan and Creelman, 2004). Accordingly, we consider
feeling generation to closely resemble sensation processes. The paradigm that we suggest provides a new perspective for the exploration of individual differences in the generation of feelings that does not rely on raw self-report or indirect correlates, which may show inconsistency in their relation to feelings (e.g., Mauss et al., 2005). Moreover, our SDT conceptualization enables us to explore feelings as composed of several distinct constituents, each with particular individual differences.

Our work concentrates on the generation of negative feeling (valence, i.e., the degree of averseness). Valence is considered to be a core dimension of the emotional experience, either as a primitive (e.g., Barrett, 2006) or a result of discrete feelings (e.g., Ekman and Cordaro, 2011). We acknowledge the arousal dimension (i.e., How intense is one’s feeling) as an equally relevant core dimension (e.g., Kron et al., 2013; Mattek et al., 2017), yet it was only controlled and not assessed in the current study.

The current study extends the application of our SDT model to probe the relation between feeling generation and several relevant constructs. We started with assessment of individual differences in five core personality traits, but we predicted that the correlations would concentrate in the affective domain. Specifically, we predicted correlations with Neuroticism given its putative close relations with Rothbart’s (2007) "negative affectivity" dimension of temperament. This temperamental dimension refers to the level of experienced fear, anger, sadness, discomfort and soothability. After establishing this unique relationship with personality, we continued to a more exploratory assessment of individual differences in the affective style domain (self-reports concerning the experience and regulation of emotion).

In the next section, we focus on the Neuroticism trait, and affective-style related constructs. Finally, we portray the building blocks of the SDT model and present tentative predictions regarding individual differences.

1.1. Neuroticism

In the Big-Five taxonomy (John and Srivastava, 1999), Neuroticism, representing a spectrum of emotional (in)stability is most strongly related to negative emotion (e.g., Watson and Clark, 1992; Robinson, 2007). Neurotic individuals experience increased emotional reactivity (Servaas et al., 2013), especially to negative triggers (Canli, 2008) and are likely to present a negative perspective on daily life. In Gray’s (1991) theory, the Behavioral Inhibition System is linked to Neuroticism (Servaas et al., 2013).

1 Neuroticism is not the only emotion-related personality trait. Past literature identified Extraversion as closely linked to positive affectivity (Robinson, 2007), which is out of the present scope.
1.2. Affective-style and difficulties

Difficulties in detecting, reporting and managing feelings are approached through the constructs of Alexithymia, and emotion regulation. Alexithymia refers to a limited ability to recognize and verbalize own emotions (Sifneos, 1996), poor recognition of bodily sensations stemming from an emotional arousal, and an external cognitive style, a concrete approach, with repetitive non-psychologically related details (Wise et al., 2004).

Emotion regulation is the attempt to influence which emotions to have, when to have them, and how these emotions are experienced and expressed (Gross et al., 2006). Previous studies link feeling generation and emotion regulation, since adaptive emotion regulation requires a clear conscious emotional experience (e.g., Barrett et al., 2001; Bonanno and Burton, 2013; Sheppes et al., 2011).

1.3. SDT model for feeling generation

In line with SDT perspective, we assume that emotional feelings rely on an inherently noisy information, analogous to that involved in sensory feelings such as seeing and hearing. Specifically, when a person is exposed to emotional signals, components of emotion become activated, including bodily and facial reactions, cognitions, action-thought tendencies and others (Scherer, 2001). The origin of the signal may be the outer world, but also an inner trigger (e.g., a disturbing thought or memory). However, at this point, our operationalization focuses on inputs from the outer world.

Following others (Barrett et al., 2007; Laird and Lacasse, 2014; Schachter and Singer, 1962; e.g., Thagard and Aubie, 2008), we assume that this information is integrated into what we term, emotional evidence. Emotional evidence is not conscious feeling, but merely the activation of the (internal) emotional system occurring in response to a signal. The model assumes the involvement of a decision process that imposes a threshold above which the emotional evidence is translated into a conscious reportable feeling (see Fig. 1). Although the emotional evidence results from the signal, other sources, from both the inner and outer world contribute to the emotional evidence making it essentially noisy. These include mood, reactions to psychoactive substances, hormones, etc. Given this noise, feeling-errors may occur: False Alarms (FA) are over-reactions to mild triggers, and Misses (Miss) are failures to consciously react to a strong trigger.
How SDT characterizes feeling generation? SDT discriminates between two independent parameters, sensitivity (here, \(d_a\)), and criterion (\(c\)) (Macmillan and Creelman, 2004). ii We describe \(d_a\) as “evidence-differentiation”, the ability of the emotional system to differentiate (in terms of its level of activation) between external triggers of varying intensity (equivalent to the distinction between signal and noise in classic SDT). In other words, to what extent our emotional evidence discriminates between intense and mild triggers. The second SDT parameter, \(c\), is termed “report-criterion”, the point above which one’s emotional sensation allow reporting (strong) feeling, or the point above which there is enough emotional evidence for intense feeling to appear.

We describe the theory as applied to the discrimination task that we used. This task requires to discriminate between very similar supra-threshold stimuli, unlike the (more familiar to most readers) detection SDT-task, requiring to detect a single stimulus whose intensity is very close to threshold. Though both variants, detection and

---

ii The simple and commonly used calculation of SDT parameters is based on a yes/no response, where the proportion of Hits (here, experiencing an intense trigger as intense) and False Alarms (experiencing a weak trigger as intense) serve as a basis for computing \(d'\) (sensitivity) and \(c\) (criterion). Here we used a rating procedure in which the response was continuous instead of yes vs. no. In this paradigm, there is one sensitivity measure but multiple (here, 5) criteria, each separating the response continuum into two levels, e.g., 1 vs. 5−6, 1−2 vs 3−6 etc. Estimation of SDT parameters in the rating procedure is based on model fitting using maximum-likelihood estimation. One advantage of the rating procedure is the ability to assess the appropriateness of the model as well as to assess which exact version of the SDT model is appropriate. Given that we found the unequal variance model to be appropriate, we computed \(d_a\) (and not \(d'\)), the appropriate measure in this case.
discrimination, are relevant to our model, the current study focuses on the ability to associate differential experienced of valence-intensity with mild vs. intense triggers.

1.4. Neuroticism and report-criterion setting

We consider \( c \) as a conceptual cousin of the construct of Neuroticism, and thus, the correlation between the two constructs was our main prediction. Specifically, we postulate that neurotic individuals require less emotional evidence in order to experience and report negative feelings as intense. Their Differential appraisal reflects the heightened tendency to appraise or construe events as relatively threatening or dangerous (Suls and Martin, 2005). In SDT, expectancy reflects the prior odds ratio of experiencing negative feelings for justifiable reasons. When priors are high, little evidence is needed to reach a high posterior probability (the odds given the evidence). Since \( c \) represents the amount of evidence needed to make a decision, it can be lowered when priors are high. Two elements contribute to the posterior odds: the prior odds and the actual likelihood ratio of justifiable negative feeling (e.g., caused by a trigger) relative to unjustifiable negative feeling (e.g., without a trigger or following a neutral trigger). In SDT, \( c \) is often expressed by this likelihood ratio (\( \beta \)) and is closely related conceptually and mathematically to \( c \) (Macmillan and Creelman, 2004). Indeed, White et al. (2016), found lowered \( c \) among highly anxious as compared to low anxious individuals. Assuming that anxiety is closely related to Neuroticism, this finding supports our prediction. However, the task in White et al.’s study involved a decision regarding the objective characteristics of the stimuli (words) rather than reporting subjective experience.

1.5. Individual differences in feelings — SDT perspective

The data reported here were gathered as a part of a large scale experiment (Karmon-Presser et al., 2018). Since the first paper focused on model validation, we were requested to removed most of the individual differences examination, which is reported here. Specifically, we concentrated on the Big Five personality traits (John and Srivastava, 1999), Alexithymia (Bagby et al., 1994) and on difficulties in regulating emotions (Gratz and Roemer, 2004). We included the Big Five mainly in order to test the reasonable prediction that only Neuroticism will be correlated with the SDT parameters. Here, the lack of correlation with the other dimensions of the Big Five provides discriminant validity. We chose Alexithymia and emotion-regulation, given their obvious putative relationship with emotion experience.

In the SDT task, we presented strongly or weakly negative pictures from the IAPS (Lang et al., 1997) and asked participants to rate their feelings on continuous scales. We also included a color decision task which was performed between picture presentations to minimize emotional carryover and disguising the goal of the experiment.
Our model decomposes the conscious experience to the ability to emotionally differentiate between negative triggers ($d_a$), and the general emotional tendency to experience/report intense negative feelings ($c$). We ask here, which one of the two (or both) is related to emotion-related personality traits, difficulties in emotional experience, and difficulties in emotion regulation. To distinguish the latter two from personality, we termed them “affective style”.

Very few of the SDT applications in emotion research are made on reports of subjective experience (Barbosa et al., 2016; Neufeld, 1975; Nielsen and Kaszniak, 2006; for the three exceptions). To our knowledge, the only SDT study providing basis of specific prediction in the current study is Barbosa et al. (2016), who showed reduced sensitivity ($d'$) and elevated response bias ($c$) in arousal experience, among anti-social men. This finding, along with meta-analysis indicating positive relation between anti-social tendency and Neuroticism level (Jones et al., 2011) suggests a prediction of negative correlation between evidence differentiation ($d_a$) and Neuroticism, and positive correlation between the report criterion ($c$) and Neuroticism. Note that the prediction regarding $c$, which stems from the (tentative) extrapolation from Barbosa et al. (2016) is opposite to that which we made beforehand. This prediction is based on our conceptualization of a $c$ as reflecting the general tendency to experience/report intense negative feelings.

Regarding affective-style, difficulties in making fine discriminations among emotions are seemingly reflected in e.g., TAS — difficulty identifying feelings, and DERS — difficulty in emotional clarity. Because we used a discrimination SDT task (which requires making fine discriminations), we tentatively predicted a correlation between these constructs lower $d_a$. The other dimensions of affective style (e.g., DERS — limited access to regulation strategies, and non-acceptance) seem to be only indirectly related to the experience itself, and we thus predicted much weaker relationships to the SDT parameters.

2. Methods

The detailed method is reported in Karmon-Presser et al. (2018), and here it is described only briefly.

Participants. Ninety-three participants (Mean age = 24, 67 females), undergraduate students, completed the experiment for either course credit or monetary compensation of approximately 23 USD. Sample-size was determined using G-Power (Faul et al., 2007), such that it would provide Power > .80 to detect $r = .30$ with a two-sided test.

SDT Task. Participants completed the task (Fig. 2) in two separate equally long sessions, on different days, in order to prevent fatigue. Stimuli (two sets of IAPS
pictures: strong vs. weak valence, based on the IAPS norms: \( M_{(\text{strong})} = 2.35, \ SD_{(\text{strong})} = 0.14, \ \text{Range}_{(\text{strong})} = 2.25-2.75, \ M_{(\text{weak})} = 3.35, \ SD_{(\text{weak})} = 0.15, \ \text{Range}_{(\text{weak})} = 3.05-3.55, \) on a 1–9 scale) were distributed evenly between sessions. The distance between the averaged valence of the two sets was small, in order to create ambiguity and difficulty in discrimination which is required for an SDT task. In addition, both the "strong" and the "weak" categories were relatively negative, thus, participants were asked to discriminate between two intense levels of negative feeling. Each category contained 71 pictures, relatively large trial sample, as customary in SDT analysis. The standard deviation of the valence rating within each category was maintained very low (\(~0.09 \) scale points), in order to reduce stimulus variability within the category. Valence categories were matched for arousal levels (Karmon-Presser et al., 2018).

**Procedure.** The task included a training session. General instructions were presented along with the rating bars (see also Karmon-Presser et al., 2018). After completing 12 trials (including all stimuli categories) participants completed the task-proper (149 trials, in two sessions, three blocks each). Participants were instructed to experience emotions naturally during pictures presentation.

After completing the SDT task, participants completed an emotion-regulation choice task (reported in Karmon-Presser et al., 2018) and then completed the questionnaires.

**Self-report measures.** Participants completed 3 questionnaires presented via an internet survey program. The order of the questionnaires (as presented here) was fixed for all participants in order to prevent order-related variance that would have lowered the statistical power of the correlational analyses. The following questionnaires were administered:

*Affective-style.* Toronto Alexithymia Scale (TAS, Bagby et al., 1994): The questionnaire has three sub-scales, (with total of 20 items): difficulty identifying feelings (e.g., "I am often confused about what emotion I am feeling"), difficulty describing feelings (e.g., "It is difficult for me to find the right words for my feelings"), and
externally-oriented feelings (e.g., "I prefer to just let things happen rather than to un-
derstand why they turned out that way"). Items were rated on a 1 to 5 scale. This
well-established questionnaire was found to be a strong predictor of various affect
related constructs, such as emotional intelligence (Saklofske et al., 2003) and psy-
chopathology (Frewen et al., 2008; e.g., Li et al., 2015).

Difficulties in Emotion Regulation Scale (DERS, Gratz and Roemer, 2004): The
questionnaire has six sub-scales, from which 4 were used (with total of 25 items):
non-acceptance (e.g., "When I’m upset, I become irritated with myself for feeling
that way"), lack of emotional awareness (e.g., ‘I’m (in)attentive to my feelings”),
limited access to emotion regulation strategies (e.g., "when I’m upset, my emotions
feel overwhelming"), lack of emotional clarity (e.g., "I am confused about how I
feel"). Items were rated on a 1 to 5 scale. Based on pilot results, we omitted two addi-
tional sub-scales from the original questionnaire—difficulties engaging in goal-
directed behavior, and impulse control difficulties. The literature provides evidence
for the construct validity of the measure (e.g., Fowler et al., 2014; Gratz and Roemer,
2004). Specifically, the questionnaire was found to be predictive (jointly with Anxiety Sensitivity) of anxiety symptoms (Kashdan et al., 2008).

Personality. Big-Five Inventory (BFI, John and Srivastava, 1999): The question-
naire assesses the Big-five personality traits: extraversion, agreeableness, conscien-
tiousness, neuroticism, openness to experience. The Questionnaire included 44
items, rated on a 1 to 5 scale. The Big-Five taxonomy, and its widely used assess-
ment tool, the BFI (John and Srivastava, 1999), were validated through numerous
studies in past literature. Big-Five traits as represented by the BFI questionnaire
were found to predict many affect-related constructs, such as positive\negative
affect (Robinson, 2007), anxiety and depression symptoms (Kotov et al., 2010),
and emotion regulation (Livingstone and Srivastava, 2012).

2.1. Overall procedure

The experiment was administered in three sessions, on different days. First two for
the completion of the SDT task (divided into two equal parts, approximately 80 min
each), and the third for the completion of the ER choice task (not presented here) and
the questionnaires. Participants signed an informed consent form at the beginning of
the first session and were shortly debriefed at the end of the third session.

2.2. Data analysis

In Karmon-Presser et al. (2018) we show (a) that SDT provides a superior model as
compared to viable alternative models, and (b) yielded the two performance indices
of the SDT model for each participant, d, and c. Briefly, in order to perform a rating
procedure (Macmillan and Creelman, 2004), the valence continuous scale was
divided into six equally long sections (in pixels), thus transforming participants answers to a 1–6 rating format. For the valence scale, only the negative part of the scale (which is relevant for the range of the IAPS pictures in the current study) was divided into six, with the lower part considered as ‘6’ and the upper as ‘1’. The frequency of responses within each range was then modeled separately for each participant, indicating non-significant lack of fit in 95% of the participants (the ratio expected by chance).

**Analytic strategy.** Individual differences in SDT parameters were examined in a correlational design. See Tables 4 and 5 for findings’ summary. Although the original sample included a large percentage of women (70%), the pattern of significant results was similar in the full sample and when only women were included. Accordingly, the results of the full sample are presented.

Given the tentative nature of part of the predictions, in our null hypothesis testing, we adopted a hierarchical analytic approach that controls for overall $\alpha$ inflation. Specifically, in the first step, we began with two omnibus tests, one for each construct. The observed variables were divided into two separate groups: (1) Affective-style measures derived from two questionnaires TAS and DERS. (2) Personality measures derived from the BFI questionnaire. We corrected $\alpha$ in these tests by dividing it by two, the number of omnibus tests that were performed ($0.05/2 = 0.025$). Only results that met significance in this step were considered for further exploration. Note that we did not apply any selection criteria to the regression analyses, in other words, all independent variables were forced to be included in the given model.

**Bayesian analysis.** Following growing recent criticism for the traditional null hypothesis testing method (e.g., Wetzels et al., 2011), we decided to complement this analysis with Bayesian significance testing, using Bayes Factors as computed using JASP 0.8.1.2 (JASP Team, 2017). Given that we assumed the prior odds of H0 and H1 to be equal, the Bayes Factor (BF) values that we report represent the relative odds of H1 and H0 (BF10 represents the relative odds of H1 relative to H0 and BF01 represents the relative odds of H0 relative to H1) (Jeffreys, 1961).

### 3. Results

**Means, Variance, and Reliabilities** (Tables 1, 2, and 3).

We used $c3$ (from now, "c"), as representative report-criterion, because it is the point separating weak from intense feelings on the 1–6 continuum.

**Big Five Personality Traits** (Table 2).

**Affective-style** (Table 3).

**Correlations with Report-Criterion** (c, Tables 4 and 5).
Omnibus testing was accomplished with multiple regression analyses.

**Personality.** The relation between $c$ and personality was analyzed using multiple regression as an omnibus test, with the five BFI scales as predictors of $c$, with $\alpha$ divided by 2 (.025). This regression analysis indicated a significant result, with strong support for H1 ($R^2 = .19$, $p = .002$, $BF_{10} = 10.46$).

Further examination of the unique contribution of each predictor indicates that only Neuroticism uniquely contributed to the prediction of $c$ ($\beta = -0.33$, $p = .002$, $BF_{10} = \cdots $)

---

**Table 1.** Descriptive statistics and reliability of the SDT measures (N=93) taken from Karmon-Presser et al. (2018).

| SDT valence measures | Mean  | SD    | $r_{tt}$ |
|----------------------|-------|-------|----------|
| $c$                  | -0.17 | 0.92  | .97      |
| $d_a$                | 0.76* | 0.46  | .38      |

*a different from zero (indifference), $p < .01$.

*b reliability of $c$ was evaluated by correlating the correlation between $c$ and the average value of the other 4 criteria. Reliability of $d_a$ was estimated by a procedure described in Karmon-Presser et al. (2018), Appendix 1.

---

**Table 2.** Mean, standard-deviation and reliability of the BFI sub-scales.

| BFI-subscales         | Mean   | SD    | $r_{tt}$ | $r_{tt}$ |
|-----------------------|--------|-------|----------|----------|
| Extraversion          | 3.31   | 0.63  | .88      | .80      |
| Agreeableness         | 3.79   | 0.51  | .79      | .68      |
| Concientiousness      | 3.63   | 0.64  | .82      | .73      |
| Neurotism             | 2.93   | 0.71  | .84      | .81      |
| Openness to experience| 3.55   | 0.50  | .81      | .76      |

*a English version reliability values are taken from John & Srivastava (1999, BFI).

*b Hebrew version reliability values are taken from Etzion and Laski (1998).

---

**Table 3.** Means, standard-deviations and reliability scores. - affective-style sub-scales.

| Questionnaires        | Mean   | SD    | $r_{tt}$ | $r_{tt}$ |
|-----------------------|--------|-------|----------|----------|
| TAS                   |        |       |          |          |
| Difficulty identifying feelings | 2.24   | 0.73  | .78—.81  | .80      |
| Difficulty describing feelings | 2.37   | 0.84  | .75      | .79      |
| Externally oriented feelings | 2.08   | 0.52  | .64—.66  | .68      |
| DERS                  |        |       |          |          |
| Non-acceptance        | 2.42   | 0.93  | .85      |          |
| Lack of emotional awareness | 2.30   | 0.41  | -        |          |
| Limited access to emotion-regulation strategies | 2.25   | 0.78  | .88      |          |
| Lack of emotional clarity | 1.94   | 0.64  | .84      |          |

*a English version reliability is taken from Bagby et al. (1994, TAS) and Gratz and Roemer (2004, DERS).

*b Hebrew version reliability values are taken from Taylor et al. (2003, TAS). Hebrew version Cronbach’s alpha for total score in the DERS was reported .93 (Pat-Horenczyk et al., 2015).
22.74). In other words, high levels of Neuroticism were related to a lower c (Fig. 3). Though Openness to experience showed a significant beta as well ($\beta = -0.23$, $p = .03$), Bayesian analysis indicated only anecdotal evidence for $H1(BF10 = 2.92)$, and it was not further probed. The remaining Big-Five traits indicated no meaningful contribution to the prediction of $c$, although in nearly all cases, the results could not support $H0$ either (Table 4).

The Bayesian regression analysis indicates that although the rest of the personality traits did not uniquely contributed to the prediction of $c$, the model including all five traits was the most suitable model, indicated by the highest BF score. In addition, examination of the correlation matrix between the Big Five and $c$ suggests that the other dimensions explained $c$-irrelevant variance in Neuroticism (possibly indicating that the other dimensions acted as suppressor-variables, see Table 6 and Table 7 for correlation matrix among the Big five showing that 3 of the four...
correlations with Neuroticism were significant, albeit quite moderate). This has led to lower correlation of Neuroticism with c (r = .23, BF\textsubscript{10} = 1.55) as compared to its substantial unique contribution (Beta) in the multiple-regression model, a discrepancy suggesting statistical suppression.

Affective-style. The relationship between Affective-style and the report-criterion (c) was examined in a multiple regression analysis in which the subscales of the two questionnaires (i.e., TAS and DERS) served as predictors of c with α divided by 2 (.025). The multiple-regression model indicated a significant result (R\textsuperscript{2} = .15, p = .04), yet the p-value did not pass the corrected Alpha level (.025) and similarly, the Bayesian analysis showed anecdotal support for H\textsubscript{0} (with BF\textsubscript{10} = 0.30, meaning BF\textsubscript{01} = 3.33). See Table 5 for summary. Given this somewhat undecided situation, we decided to leave further exploration for future research.

3.1. Evidence-differentiation (d\textsubscript{a}) analysis

Personality. The relation between d\textsubscript{a} and personality was analyzed as before. This regression analysis indicated a null result, with strong evidence for H\textsubscript{0}, as indicated by the BF (R\textsuperscript{2} = .02, p = .82, BF\textsubscript{01} = 100.00). Simple correlations between d\textsubscript{a} and the five dimensions indicated no meaningful correlations, as indicated by the BF (Table 4). Moreover, for both Neuroticism and Openness to experience, BF indicated substantial evidence for H\textsubscript{0} (BF\textsubscript{01} = 4.00 and BF\textsubscript{01} = 4.17, respectively). Thus, the entire personality domain (as represented through the Big-Five) was found to be unrelated to differences in evidence-differentiation.

![Fig. 3. Scatterplot of the significant correlation between c and Neuroticism (r = -.23, BF\textsubscript{10} = 1.55).](https://doi.org/10.1016/j.heliyon.2019.e01344)
Affective-style. Relationship between Affective-style and $d_a$ was examined as before. The multiple-regression model indicated a significant result, with very strong evidence for H1 ($R^2 = .27$, $p < .001$, $BF_{10} = 70.37$), such that affective-style predicted evidence-differentiation levels. Next, we examined both the unique contribution of each subscale (Beta) and the correlation matrix between all subscales and $d_a$. Only subscales showing meaningful results in both aspects (based on Bayesian analysis) were considered consistent and reported here. See Table 5.

In DERS, the subscale, "limited access to emotion regulation strategies" (being overwhelmed by your own negative feeling, with difficulties to provide means to change them and recover from them) was a meaningful predictor of $d_a$ ($\beta = -0.61$, $BF_{10}=501.25$). Thus, limited access to regulation strategies was related to low $d_a$. Bayesian analysis of the correlation matrix between the questionnaires and $d_a$ indicated substantial support for the correlation ($r = -.29$, $BF_{10} = 7.30$). This correlation shows that being overwhelmed by your own negative feeling, with difficulties to provide means to change them and recover from them is related to poor ability to emotionally differentiate between negative triggers of different intensities.

Note that the DERS “non-acceptance” subscale was also a meaningful predictor of $d_a$. However, this contribution was not supported by a meaningful simple correlation between the two constructs. The high correlation between “non-acceptance” and “limited access to regulation strategies” may explain this gap, with the latter seemingly acting as a suppressor variable in the regression model. Thus, this subscale was not further examined.

3.2. Five criteria analysis

Using a discrimination task and not a detection task left two equally plausible interpretations for $c$ differences. First, these individual differences may represent a shift in the report criterion. For example, low levels of Neuroticism may reflect a bias toward reporting, and consciously feeling, intense negative emotions (henceforth, “genuine
Table 7. Correlational matrix of TAS (Alexithymia) and DERS (Regulation difficulties) (n=93).

|                          | correct_non-acceptance | correct_lack of emotional awareness | correct_limited access to e.regulation strategies | correct_lack of emotional clarity | correct_Difficulty identifying feelings | correct_Difficulty describing feelings | correct_Externally oriented feelings |
|--------------------------|------------------------|------------------------------------|--------------------------------------------------|----------------------------------|----------------------------------------|--------------------------------------|-------------------------------------|
| correct_non-acceptance   | —                      |                                    |                                                  |                                  |                                        |                                      |                                     |
| correct_lack of emotional awareness | -0.019                |                                    |                                                  |                                  |                                        |                                      |                                     |
| correct_limited access to e.regulation strategies | 0.632 | 0.042 | — |                                  |                                        |                                      |                                     |
| correct_lack of emotional clarity | 0.509***               | 0.271                              | 0.536***                                         | —                                |                                        |                                      |                                     |
| correct_Difficulty identifying feelings | 0.483***               | 0.201                              | 0.542***                                         | 0.699***                         |                                        |                                      |                                     |
| correct_Difficulty describing feelings | 0.126***               | 0.443***                           | 0.305*                                           | 0.538***                         | 0.466                                  |                                      |                                     |
| correct_Externally oriented feelings | -0.033                | 0.474***                           | 0.156                                            | 0.185                            | 0.190***                               | 0.412***                            | —                                   |
criterion shift”). According to the second interpretation, different people require similar amounts of emotional evidence in order to experience intense negative feeling (i.e., there is no "genuine criterion shift"). However, they have different baseline levels of emotional evidence. In accordance, highly neurotic people actually "feel more", or have higher initial levels of emotion, such that their emotional evidence distribution is shifted upwards. Therefore, in this case we would expect equal report criterion for different levels of Neuroticism, but different absolute emotional sensation (henceforth, “baseline emotional evidence differences”). The two interpretations can also coexist.

In order to shed light on the source of difference in \( c \), we adopted the method first used in Karmon-Presser, Harpaz & Meiran (under review). This method capitalizes on the fact that we used a rating procedure in which participants provide a rating instead of a yes/no answer. SDT views rating as reflecting multiple criteria. For example, the reason to rate a trigger as belong to Intensity = 2 and not to Intensity = 1 implies that there is a criterion placed between Intensity 1 and 2. Since we grouped the ratings into six categories, this implies that we had five criteria. In the ANOVA, Neuroticism was a between-subjects independent variable, which was formed by dividing the continuous Neuroticism score into three levels (high, moderate, low). Another independent variable was Criterion (c1 to c5) and the dependent variable was criterion level (see Fig. 4). In this analysis, the critical

\[
\text{Fig. 4. Scatterplots of the significant correlation between } d_a \text{ and DERS — Limited access to emotion-regulation strategies. Correlation was re-computed excluding the outlier and remained significant, } (r = -.28, p = .006, BF_{10} = 5.17).
\]
finding is the interaction between Criterion and Neuroticism. This interaction tests whether Neuroticism has a differential effect on the five criteria. Such a differential effect, if exists, represents the involvement of a genuine criterion shift for high neurotic people compared to low neurotic. If the correlation between c3 and Neuroticism reflects baseline emotional evidence differences only, there should not be an interaction. The reason is that if the criteria remain fixed, it is just that the emotional evidence of high-Neuroticism participants is shifted upward. In contrast, a significant interaction indicates that a genuine criterion shift must also play a role.

The results of a mixed-model B/ANOVA indicated substantial evidence for an interaction between Neuroticism and criterion (F(8,364) = 2.91, p = .004, ηp² = .06 BF₁₀ = 4.86) see Fig. 5. Thus, the pattern of results is at least partially explained by a genuine downward criterion shift among high neurotic people.

4. Discussion

The current study employs a new SDT framework for assessment of feeling, which we consider as governed by two parameters. Results indicate a double dissociation, as c but not da was related to Neuroticism, yet da (but not c) was related to self-reporting difficulties to successfully apply regulation strategies. This pattern of findings further validates the SDT parameters, as representing distinct components of feeling, not only theoretically but also in individual differences.

Among the Big-Five, the report-criterion, c, was found to be correlated only with Neuroticism, such that high levels of Neuroticism were related to a lower c. The presence of a report-criterion involvement was further validated using the five criteria analysis Interestingly, the prediction derived from anti-social behavior studies (Barbosa et al., 2016; Jones et al., 2011) was not supported, as Neuroticism was accompanied with lowering of c. This difference may be due to the fact that
lower-order Neuroticism traits (mainly Impulsivity and Angry Hostility) contributed to the positive relation with anti-social behavior (Jones et al., 2011). However, we did not assess these lower-order traits of the Big Five in the current study. Tentatively, we assume that facet scales such as Impulsivity and Hostility are less related, as compared with Anxiety and Vulnerability.

To our understanding, the report-criterion, as a trait, represents the individual baseline of conscious negative affectivity, which is a central aspect of Neuroticism. This finding accords with past literature which suggests that Neuroticism is systematically related to a tendency to experience more negative feelings in reaction to negative events (Watson and Clark, 1992), and also Robinson’s conceptualization (2007) of Neuroticism as affective memory network, in which negative affect is favored. It also accords a similar finding by White et al. (2016) regarding preferential processing of threatening information in anxiety.

While $c$ and Neuroticism are correlated, these constructs are not identical. First, $c$ may be considered as reflecting both trait and state. Specifically, the traditional SDT treatment of the criterion ($c$) is considered as the response bias, a parameter changing as a function of context (e.g., the cost of misses vs. false alarms). Thus, a-priori, $c$ is considered in the SDT literature as being related to a state. In the current work, however, we provide a somewhat counterintuitive demonstration of $c$ as a trait. Here, $c$ may be described as the emotional tendency for responding to the world. This could be, for example, the tendency of a neurotic person to view the world as bad or dangerous.

Indeed, there are aspects of neuroticism that are closely related to $c$ (i.e. the heightened emotional baseline towards negative context, Robinson, 2007). However, one would also expect a-priori that neuroticism will be correlated with $d_a$. Specifically, a plausible (yet somewhat complex) hypothesis is that Neuroticism is characterized by poor $d_a$, and that the lowered criterion is adopted in order to compensate for the potential resultant increase in Miss rate. We note that this speculation was not supported by our data. In any event, the fact that we found a correlation only with $c$, further indicates the distinction between the constructs, along with the novel information portrayed through decomposing feeling based on SDT parameters. Finally, Neuroticism covers additional aspects that are seemingly less related to $c$, such as emotional instability (Ormel et al., 2012).

Regulatory Focus Theory (Crowe and Higgins, 1997) distinguishes between promotion-focus and prevention-focus. While promotion-focus leads to exploratory seeking of gain, prevention-focus leads to an attempt to minimize losses. Interestingly, promotion and prevention foci were conceptualized in SDT terms as eagerness, or behavior which maximize "Hits" in the first, and vigilance, or behavior that minimize "False-Alerts" in the latter (Crowe and Higgins, 1997). In other words, the different foci reflect different $c$ setting for behavior. Naturally,
Neuroticism is linked to prevention-focus (Manczak et al., 2014). While being similar to these theories, our theory focuses on feelings whereas they focus on actual behaviors. Given that \( c \) is determined, in part, by the relative costs of Hits and False-Alarms, this distinction between feeling and behavior is an important one. Specifically, the costs of wrong avoidance behaviors are quite different from the costs of expressing negative feelings.

Notably, in the current study, \( c \) was not significantly related to aspects of emotion regulation style as represented by the DERS. This lack of correlation was counter to our initial predictions, as we consider \( c \) as being relevant to decisions regarding emotion regulation. For example, in a previous study (Karmon-Presser et al., 2018) we showed a relation between \( c \) and emotion regulation choice, such that low \( c \) was correlated with a tendency to prefer distraction over reappraisal when experiencing intense negative feelings. Given the discrepant findings, we suggest that the DERS may reflect the reported (retrospective) experience of regulating, while our past findings reflect actual trial-by-trial regulation choices.

Contrary to \( c \), \( d_a \), was not correlated to any Big-Five personality trait. The fact that the level of (negative valence) differentiation is not an inherent feature of Neuroticism is less intuitive to us. In retrospect we suggest that it may reflect the nature of Neuroticism, which to our understanding is connected to one’s interaction with the world and his/her expectations from it (reflected by \( c \)), and is less connected to the ability to gather and integrate the emotional evidence differentially, according to trigger intensity. These aspects of the process might occur at a sub-conscious level that is unrelated to Neuroticism.

In the Affective-style domain, \( d_o \) was related to a subscale in the DERS questionnaire such that low \( d_o \), representing lower emotional differentiation, was related to greater experienced difficulty in overcoming negative feelings. This experience was assessed retrospectively based on statements such as: "When I’m upset, my emotions feel overwhelming" which are included in the scale "limited access to emotion regulation strategies".

Since we did not predict this finding, we can only offer some post-hoc explanations. First, it is plausible, that choosing an appropriate regulation strategy requires reliable knowledge regarding one’s inner state, and when such knowledge is lacking, the chosen strategies are sub-optimal and thus relatively ineffective. The core ability to emotionally differentiate between different intensities is crucial in later processes of applying a specific regulation strategy, in line with past literature, suggesting that clear perception of one’s feelings (in terms of intensity), is important for adaptive emotion regulation (Sheppes et al., 2011). An alternative approach may suggest that efficient regulation requires continuous internal feedback. It may thus be that people with low \( d_o \) do try to regulate, yet the relevant internal evidence is not
properly accumulated, and thus, fail to experience the change in feeling and remain "overwhelmed".

Clearly, the current study suffers from several limitations. The first is the relatively small sample size for a correlational study. Second, the chosen questionnaires for assessment of both general personality traits and affective-style are only "the tip of the iceberg" when considering relevant constructs. Our analysis also did not assess gender differences. Moreover, measures that are not based on self-report are lacking and should be incorporated in future studies since they may provide another angle for understanding the SDT model. Note also that the current analysis was limited to non-clinical population. To summarize, our model assumes that feelings can be erroneous, as people may overreact or underreact to emotional triggers. Accordingly, people differ in the amount of these errors (reflected in $d_a$), the evidence-differentiation, but also in their tendency toward a specific kind of error (Misses - or under-reaction, vs. False Alarms, or overreaction), as reflected in their report-criterion, $c$. Our findings suggest that affect related personality trait (Neuroticism) is related to one’s unique pattern of errors in feeling generation (the location of the report-criterion, indicating preference for False-Alarms) but not to the level of differentiation. However, the accessibility of various regulation strategies was related to the emotional evidence-differentiation, but not the report-criterion.

**Declarations**

**Author contribution statement**

Anat Karmon-Presser, Nachshon Meiran: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

**Funding statement**

Nachshon Meiran was supported by Israel Science Foundation (#381 \ 15).

**Competing interest statement**

The authors declare no conflict of interest.

**Additional information**

No additional information is available for this paper.
References

Pat-Horenczyk, R., Cohen, S., Ziv, Y., Achituv, M., Asulin-Peretz, L., Blanchard, T.R., et al., 2015. Emotion regulation in mothers and young children faced with trauma. Infant Ment. Health J. 36 (3), 337–348.

Bagby, R.M., Parker, J.D., Taylor, G.J., 1994. The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. J. Psychosom. Res. 38, 23–32.

Barbosa, F., Almeida, P.R., Ferreira-Santos, F., Marques-Teixeira, J., 2016. Using signal detection theory in the analysis of emotional sensitivity of male recidivist offenders. Crim. Behav. Ment. Health 26, 18–29.

Barr, L.K., Kahn, J.H., Schneider, W.J., 2008. Individual differences in emotion expression: hierarchical structure and relations with psychological distress. J. Soc. Clin. Psychol. 27, 1045–1077.

Barrett, L.F., 2006. Valence is a basic building block of emotional life. J. Res. Personal. 40, 35–55.

Barrett, L.F., Gross, J., Christensen, T.C., Benvenuto, M., 2001. Knowing what you’re feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation. Cognit. Emot. 15, 713–724.

Barrett, L.F., Mesquita, B., Ochsner, K.N., Gross, J.J., 2007. The experience of emotion. Annu. Rev. Psychol. 58, 373–403.

Bonanno, G.A., Burton, C.L., 2013. Regulatory flexibility: an individual differences perspective on coping and emotion regulation. Perspect. Psychol. Sci. 8, 591–612.

Bradley, M.M., Lang, P.J., Cuthbert, B.N., 1993. Emotion, novelty, and the startle reflex: habituation in humans. Behav. Neurosci. 107, 970–980.

Bradley, M.M., Cuthbert, B.N., Lang, P.J., 1996. Picture media and emotion: effects of a sustained affective context. Psychophysiology 33, 662–670.

Canli, T., 2008. Toward a neurogenetic theory of neuroticism. Ann. N. Y. Acad. Sci. 1129, 153–174.

Crowe, E., Higgins, E.T., 1997. Regulatory focus and strategic inclinations: promotion and prevention in decision-making. Organ. Behav. Hum. Decis. Process. 69, 117–132.

Ekman, P., Cordaro, D., 2011. What is meant by calling emotions basic. Emot. Rev. 3, 364–370.
Etzion, D., Laski, S., 1998. Big five questionnaire (Hebrew version by permission). Tel Aviv University, Faculty of Management, the Institute of Business Research, Tel Aviv, Israel.

Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39, 175–191.

Fowler, J.C., Charak, R., Elhai, J.D., Allen, J.G., Frueh, B.C., Oldham, J.M., 2014. Construct validity and factor structure of the difficulties in Emotion Regulation Scale among adults with severe mental illness. J. Psychiatr. Res. 58, 175–180.

Frewen, P.A., Dozois, D.J., Neufeld, R.W., Lanius, R.A., 2008. Meta-analysis of alexithymia in posttraumatic stress disorder. J. Trauma. Stress, 21. Official Publication of The International Society for Traumatic Stress Studies, pp. 243–246.

Gratz, K.L., Roemer, L., 2004. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. J. Psychopathol. Behav. Assess. 26, 41–54.

Gray, J.A., 1991. Neural systems, emotion, and personality. In: Madden IV, J. (Ed.), Neurobiology of Learning, Emotion, and Affect. Raven Press, New York, pp. 273–306.

Gross, J.J., Richards, J.M., John, O.P., 2006. Emotion regulation in everyday life. In: Snyder, D.K., Simpson, J., Hughes, J.N. (Eds.), Emotion Regulation in Couples and Families: Pathways to Dysfunction and Health. American Psychological Association, Washington, DC, US, pp. 13–35.

JASP Team, 2017. JASP [Computer software], 0.8.1.2.

Jeffreys, S.H., 1961. Theory of Probability. Clarendon Press, Oxford.

John, O.P., Srivastava, S., 1999. The Big Five trait taxonomy: history, measurement, and theoretical perspectives. Handb. Pers. Theory Res. 2, 102–138.

Jones, S.E., Miller, J.D., Lynam, D.R., 2011. Personality, antisocial behavior, and aggression: a meta-analytic review. J. Crim. Justice 39, 329–337.

Karmon-Presser, A., Sheppes, G., Meiran, N., 2018. How does it “feel”? A signal detection approach to feeling generation. Emotion 18, 94–115.

Kashdan, T.B., Zvolensky, M.J., McLeish, A.C., 2008. Anxiety sensitivity and affect regulatory strategies: individual and interactive risk factors for anxiety-related symptoms. J. Anxiety Disord. 22, 429–440.
Kotov, R., Gamez, W., Schmidt, F., Watson, D., 2010. Linking “big” personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. Psychol. Bull. 136, 768–821.

Kron, A., Goldstein, A., Lee, D.H.-J., Gardhouse, K., Anderson, A.K., 2013. How are you feeling? Revisiting the quantification of emotional qualia. Psychol. Sci. 24, 1503–1511.

Laird, J.D., Lacasse, K., 2014. Bodily influences on emotional feelings: accumulating evidence and extensions of William James’s theory of emotion. Emotion Review 6, 27–34.

Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1997. International Affective Picture System (IAPS): Technical Manual and Affective Ratings. NIMH Center for the Study of Emotion and Attention, pp. 39–58.

Li, S., Zhang, B., Guo, Y., Zhang, J., 2015. The association between alexithymia as assessed by the 20-item Toronto Alexithymia Scale and depression: a meta-analysis. Psychiatr. Res. 227, 1–9.

Livingstone, K.M., Srivastava, S., 2012. Up-regulating positive emotions in everyday life: strategies, individual differences, and associations with positive emotion and well-being. J. Res. Personal. 46, 504–516.

Macmillan, N.A., Creelman, C.D., 2004. Detection Theory: A User’s Guide. Psychology press.

Manczak, E.M., Zapata-Gietl, C., McAdams, D.P., 2014. Regulatory focus in the life story: prevention and promotion as expressed in three layers of personality. J. Personal. Soc. Psychol. 106, 169–181.

Mattek, A.M., Wolford, G.L., Whalen, P.J., 2017. A mathematical model captures the structure of subjective affect. Perspect. Psychol. Sci. 12, 508–526.

Mauss, I.B., Levenson, R.W., McCarter, L., Wilhelm, F.H., Gross, J.J., 2005. The tie that binds? Coherence among emotion experience, behavior, and physiology. Emotion 5, 175–190.

Neufeld, R.W., 1975. Effect of cognitive appraisal on d’ and response bias to experimental stress. J. Personal. Soc. Psychol. 31, 735–743.

Nielsen, L., Kaszniaik, A.W., 2006. Awareness of subtle emotional feelings: a comparison of long-term meditators and nonmeditators. Emotion 6, 392–405.

Ormel, J., Riese, H., Rosmalen, J.G., 2012. Interpreting neuroticism scores across the adult life course: immutable or experience-dependent set points of negative affect? Clin. Psychol. Rev. 32, 71–79.
Robinson, M.D., 2007. Personality, affective processing, and self-regulation: toward process-based views of extraversion, neuroticism, and agreeableness. Social and Personality Psychology Compass 1, 223–235.

Rothbart, M.K., 2007. Temperament, development, and personality. Curr. Dir. Psychol. Sci. 16, 207–212.

Saklofske, D.H., Austin, E.J., Minski, P.S., 2003. Factor structure and validity of a trait emotional intelligence measure. Pers. Indiv. Differ. 34, 707–721.

Schachter, S., Singer, J., 1962. Cognitive, social, and physiological determinants of emotional state. Psychol. Rev. 69, 379–399.

Scherer, K.R., 2001. Appraisal considered as a process of multilevel sequential checking. In: Sherer, K.R., Schorr, A., Johnstone, T. (Eds.), Appraisal Processes in Emotion: Theory, Methods, Research. Oxford University Press, Oxford, pp. 92–120.

Servaas, M.N., van der Velde, J., Costafreda, S.G., Horton, P., Ormel, J., Riese, H., Aleman, A., 2013. Neuroticism and the brain: a quantitative meta-analysis of neuroimaging studies investigating emotion processing. Neurosci. Biobehav. Rev. 37, 1518–1529.

Sheppes, G., Scheibe, S., Suri, G., Gross, J.J., 2011. Emotion-regulation choice. Psychol. Sci. 22, 1391–1396.

Sifneos, P.E., 1996. Alexithymia: past and present. Am. J. Psychiatry 153, 137–142.

Suls, J., Martin, R., 2005. The daily life of the garden-variety neurotic: reactivity, stressor exposure, mood spillover, and maladaptive coping. J. Personal. 73, 1485–1510.

Taylor, G.J., Bagby, R.M., Parker, J.D., 2003. The 20-Item Toronto Alexithymia Scale: IV. Reliability and factorial validity in different languages and cultures. J. Psychosom. Res. 55, 277–283.

Thagard, P., Aubie, B., 2008. Emotional consciousness: a neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience. Conscious. Cognit. 17, 811–834.

Watson, D., Clark, L.A., 1992. On traits and temperament: general and specific factors of emotional experience and their relation to the five-factor model. J. Personal. 60, 441–476.

Wetzels, R., Matzke, D., Lee, M.D., Rouder, J.N., Iverson, G.J., Wagenmakers, E.-J., 2011. Statistical evidence in experimental psychology: an empirical comparison using 855 t tests. Perspect. Psychol. Sci. 6, 291–298.
White, C.N., Skokin, K., Carlos, B., Weaver, A., 2016. Using decision models to decompose anxiety-related bias in threat classification. Emotion 16, 196–207.

Wise, T.N., Kheriaty, A.D., Sheridan, M.J., 2004. Attribution of cause by patients with depression, anxiety, and alexithymia. Psychol. Rep. 94, 259–263.