



Published in final edited form as:

Psychol Assess. 2018 November ; 30(11): 1454–1467. doi:10.1037/pas0000584.

The Perceived Invalidation of Emotion Scale (PIES): Development and Psychometric Properties of a Novel Measure of Current Emotion Invalidation

Melissa J. Zielinski^{a,b,*} and Jennifer C. Veilleux^b

^aUniversity of Arkansas for Medical Sciences, Psychiatric Research Institute, Brain Imaging Research Center, Little Rock, AR 72205

^bUniversity of Arkansas, Department of Psychological Science, Fayetteville, AR 72701 USA

Abstract

Emotion invalidation is theoretically and empirically associated with mental and physical health problems. However, existing measures of invalidation focus on past (e.g., childhood) invalidation and/or do not specifically emphasize invalidation of emotion. In this paper, we articulate a clarified operational definition of emotion invalidation, and use that definition as the foundation for development of a new measure of current perceived emotion invalidation across a series of five studies. Study 1 was a qualitative investigation of people's experiences with emotional invalidation from which we generated items. An initial item pool was vetted by expert reviewers in Study 2 and examined via exploratory factor analysis in Study 3 within both college student and online samples. The scale was reduced to 10 items via confirmatory factor analysis in Study 4, resulting in a brief but psychometrically promising measure, the Perceived Invalidation of Emotion Scale (PIES). A short-term longitudinal investigation (Study 5) revealed that PIES scores had strong test-retest reliability, and that greater perceived emotion invalidation was associated with greater emotion dysregulation, borderline features and symptoms of emotional distress. In addition, the PIES predicted changes in relational health and psychological health over a one-month period. The current set of studies thus presents a psychometrically promising and practical measure of perceived emotion invalidation that can provide a foundation for future research in this burgeoning area.

Keywords

Emotion invalidation; Invalidation; Emotion; Self-Report; Social Relationships; Perception; Borderline Personality Disorder; Biosocial Theory; Pain

*Corresponding Author: Melissa J. Zielinski, Ph.D., Brain Imaging Research Center, Psychiatric Research Institute, University of Arkansas for Medical Sciences, 4301 W. Markham Street Little Rock, AR, 72205, MJZielinski@uams.edu, P: 501-526-8341.

Financial Disclosures: No authors reported biomedical financial interests or other potential conflicts of interest.

Introduction

Social sharing of emotions is ubiquitous across cultures, and even hardwired into human biology (Ekman & Friesen, 1971; Elfenbein & Ambady, 2002). Beyond serving as internal signals, emotions have an interpersonal function –namely, communicating needs and desires to others (Shariff & Tracy, 2011). However, others are not always responsive to sharers' expressed emotions; they may respond with negative, emotionally invalidating reactions (e.g., dismissing, criticizing, ignoring) that may exert deleterious consequences on physical and mental health. For example, invalidation is theorized to contribute to the development of borderline personality disorder (Crowell, Beauchaine, & Linehan, 2009; Linehan, 1993), and is associated with chronic pain (Linton, Boersma, Vangronsveld, & Fruzzetti, 2012), eating disorders (Haslam, Arcelus, Farrow, & Meyer, 2012; Haslam, Mountford, Meyer, & Waller, 2008; Mountford, Corstorphine, Tomlinson, & Waller, 2007), rheumatic diseases (Cano, Leong, Williams, May, & Lutz, 2012; Kool et al., 2010; Kool, Middendorp, Lumley, Bijlsma, & Greenen, 2013; Kool & Geenen, 2012), and serious mental illness (Sells, Black, Davidson, & Rowe, 2008). These studies suggest that emotion dysregulation, which is characteristic of many psychological disorders, may stem from and/or be influenced by how social partners respond to an individual's expressed emotions. Given the millions of people currently suffering from emotional disorders (Center for Behavioral Health Statistics, 2015), understanding modifiable factors that contribute to emotion dysregulation, such as emotion invalidation, is critical.

Defining Emotion Invalidation

In the present investigation, we define emotion invalidation as any social exchange during which an individual's expressed emotions or affective experiences are met with a response from another person that is perceived by the individual as implying that their emotions or affective experiences are incorrect or inappropriate. This definition characterizes emotion invalidation as an active process occurring in response to affective communication, and conspicuously prioritizes the perception of the individual sharer over the intent or objective behavior of the respondent (Ford, Waller, & Mountford, 2011; Linehan, 1997). This critical conceptual choice – to focus on perceptions – was made because appraisals causally influence emotional responses and salience during interpersonal interactions, contribute to psychological disorder, and may be valuable *intrapersonal* targets for intervention.

Measuring Emotion Invalidation

Of the three existing self-report measures of invalidation, perhaps the most popular is the Invalidating Childhood Environment Scale (ICES) (Mountford et al., 2007), which was initially developed for use with eating disorder patients. However, ICES items are framed to focus on parental behavior (rather than how this behavior was experienced by the respondent) and do not specifically assess invalidation of *emotion*. Research on the psychometric properties of the ICES in nonclinical samples has also been mixed (Mountford et al., 2007; Robertson, Kimbrel, & Nelson-Gray, 2013). Therefore, while the ICES can be used to retrospectively assess childhood invalidation, it has multiple limitations.

A second measure, the Illness Invalidation Inventory (I*3) (Kool et al., 2010; Kool, van Middendorp, Boeije, & Geenen, 2009), was developed to assess invalidation in rheumatic disease patients and has demonstrated strong psychometric properties. The I*3 is grounded in patient experiences and the authors clearly described the measure development process. However, the I*3 contains items that inquire specifically about invalidation of medical conditions and is thus not generalizable to other populations.

One final measure, the Socialization of Emotions Scale (SES) (Krause, Mendelson, & Lynch, 2003; Sauer & Baer, 2010), can be used to assess childhood emotion invalidation. The SES has good psychometric properties and comprises two factors (validation and invalidation) that index responses to emotion. However, the SES is not possible to adapt for use with adults as it contains scenario-based descriptions of parental responses to childhood emotion expression.

Developing a Novel Measure of Perceived Emotion Invalidation

Our investigation comprised a series of five studies in which we developed and assessed the psychometric properties of the Perceived Invalidation of Emotion Scale (PIES), a novel measure of emotion invalidation. Notably, we sought to examine current levels of emotion invalidation, given concerns about the accuracy of retrospective reporting and emerging evidence that current emotion invalidation is linked to health outcomes (Leong, Cano, & Johansen, 2011; Vangronsveld & Linton, 2012). Consistent with our operationalization of emotion invalidation, we placed our focus on individuals' perceptions of emotion invalidation, rather than behavioral indicators of emotion invalidation meant to be rated by an observer.

Study 1 took a qualitative approach to generating themes relevant to invalidation. These themes were then used to generate measure items. Studies 2–5, which focused on scale construction and validation, were based on established scale design guidelines (Clark & Watson, 1995; Gehlbach & Brinkworth, 2011; Reise, Waller, & Comrey, 2000). Study 2 used expert review to assess and select items for inclusion in the initial PIES item pool. An exploratory factor analysis of the selected items was conducted in Study 3, which also assessed internal consistency and convergent validity. A confirmatory factor analysis of the PIES was then conducted in Study 4. Lastly, Study 5 involved a short-term longitudinal examination of the predictive validity, incremental validity, and test-retest reliability of the finalized PIES measure. Convergent and divergent validity were also assessed. Table 1 presents demographic information about the participants in each study. All procedures were approved by the Institutional Review Board at the University of Arkansas.

Study 1 – Qualitative Inquiry

To maximize the likelihood of creating PIES items that mirror human experiences, we began our investigation with a qualitative study that examined how people experience and describe emotion invalidation. As is common in rigorous qualitative research, we used multiple methods (individual interviews and focus groups) to reduce the risk that conclusions drawn from the data would reflect systematic biases based on method.

Participants and Procedure

Twenty-two adults (ages 18–69 years) completed Study 1, which comprised individual essay questions and focus group sessions with up to six participants. They were initially escorted to a computer laboratory and told that the purpose of the study was to understand experiences in interpersonal relationships. After consenting, they completed digitized essay questions before moving to a larger space for the focus groups. All participants were undergraduate students ($n = 12$; $M_{\text{age}} = 19.33$) or community adults recruited via electronic ads ($n = 10$; $M_{\text{age}} = 40.30$). Students received course credit and community adults received 25 US dollars as compensation.

Narrative queries—All narrative queries included consideration of valence (i.e., shared emotions that were positive, negative, or neutral) and intensity (i.e., implications that shared emotions were too intense or not intense enough). The individual essay questions asked participants about others' responses to their emotions, and specifically prompted for expected responses and unpleasant responses to shared emotions, as well as responses to participants' lack of emotion during situations that emotional expression was socially expected. The focus group questions fostered more elaborative discussion about the individual essay prompts, provided an opportunity for participants to describe responses to shared emotions that they had witnessed but not experienced, and finally defined emotion invalidation and asked participants to share times in which they had experienced invalidation in response to positive, negative, and absent emotions. All Study 1 questions are available in our online supplemental material.

Qualitative coding—Participants' responses were examined using the descriptive coding method (Saldaña, 2013) following verbatim transcription and de-identification of verbal responses. Coders (MJZ and JCV and two undergraduate research assistants trained in qualitative methods), individually identified and coded all social responses to emotions described by participants. The independently assigned codes were discussed as a team and codes that were negatively valenced were compiled into a codebook with standardized phrasing (which was agreed upon through discussion) and representative examples.¹ Coders then met with the authors' larger research team and collaboratively reduced the negative codes (which were the only codes examined since emotion invalidation is inherently negative) to those responses that were agreed to have potential overlap with emotion invalidation.

Results and Discussion

We generated 19 descriptive codes thought to overlap with emotion invalidation, including: (1) *direct invalidation of emotion* (i.e., responses that clearly identify an emotion or affective experience and construe it as invalid), (2) *broad invalidation* (i.e., responses that summarize an emotional response set, without identifying a specific emotion, and construe it as invalid),

¹For example, one coder initially referred to examples in the not mirror/match emotions code as “lack of matched response” while the other two referred to these examples as “not match emotions” and “unmatched emotions.” Several participants directly used the word “mirror” when describing these experiences (e.g., “I honestly can't think of a single time when I didn't mirror someone's emotional experience of something as they've related [sic] it to me”). The standardized phrasing that was agreed upon (i.e., not mirror/match emotions) combined these concepts and participant language.

(3) *invalidation by group membership* (i.e., responses that imply that what the individual is feeling is inappropriate based on personal characteristics, such as gender, religion, or political preferences), (4) *criticize emotional response* (i.e., responses that question or critique an individual's emotional response/set of responses), (5) *general demeaning response* (i.e., responses that are attacking, directly or indirectly), (6) *get upset* (i.e., express negative personal reactions at another's shared emotions), (7) *not take seriously* (i.e., responses such as laughing or joking at another's emotions), (8) *disregard my feelings* (i.e., responses that are perceived as setting aside the individual's shared emotions), (9) *tell me how I should feel* (i.e., responses that direct the individual to feel a particular emotion/affective experience), (10) *try to change my emotions* (i.e., responses that attempt to increase, decrease, or shift the individual's emotional response), (11) *question my emotions* (i.e., responses that identify and question the individual's emotions), (12) *overreact* (i.e., expressed personal reactions that exceed the intensity of the individual's own emotions), (13) *not mirror/match emotions* (i.e., responses that involve a lack of expected shared emotional experience), (14) *not understand me* (i.e., responses that communicate lack of comprehension of the individual's emotion/affective experience), (15) *not take my side* (i.e., responses that communicate agreement with an emotional experience other than the individual's), (16) *indifference* (i.e., responses that communicate failure to care about individual's emotional experience, including complete absence of a response), (17) *sterile response* (i.e., responses that minimally acknowledge the individual's emotion/affective experience), (18) *actively avoid conversation* (i.e., responses that intentionally dissuade further discussion after an emotion/affective experience has been shared), and (19) *change the topic* (i.e., responses that move the focus away from an individual's expressed emotion). Representative examples of each code are available in our online supplemental material.

Some thematic codes, such as direct invalidation of emotion and broad invalidation of emotion (responses such as “Don't be upset, you have no reason to be upset.” and “You should get over it,” respectively), were expected to emerge based on prior research and theory, and represented prototypical invalidating experiences (Linehan, 1993; Reeves, James, Pizzarello, & Taylor, 2010; Woodberry, Gallo, & Nock, 2008). The emergence of other codes, such as not mirror/match emotions (responses like “That's not what you want to hear...you want people to be happy when you're happy and sad when you're sad.”), was more surprising and novel. These codes highlight the strength of beginning measure design with qualitative inquiry rather than solely consulting existing empirical literature. Together, the constellation of emotion invalidation varied widely and included responses that ranged in terms of intensity, activity, and degree of negative valence. However, the differences between discrete codes were relatively nuanced. Therefore, as we moved into item generation, we expected emotion invalidation to be unidimensional and aimed to measure it as such.

Initial Item Pool Construction

Approach to Item Generation—Items for the PIES were constructed from the descriptive codes and examples identified in Study 1, with an emphasis on retaining participants' own language for describing emotion invalidation whenever possible. Item phrasing also emphasized responses to shared *emotions*. Every descriptive code was represented by at least one item, and many codes were conceptually related to more than one

item. The item pool was designed to be over-inclusive and consistent with standard practice in both item construction and scale development recommendations (Clark & Watson, 1995; Gehlbach & Brinkworth, 2011).

Item Content and Anchors—There were 37 items included in the initial item pool (see supplemental material). Instructions requested that respondents reflect on their experiences with how others have responded to their emotions during the past month. Items were rated on a 5-point Likert scale ranging from 1 (Almost never; 0–10%) to 5 (Almost always; 91–100%).²

Study 2 – Expert Review

Study 2 was an expert review of the initial PIES item pool by four experts in emotion invalidation (two external and two internal). The two external experts, who were uninvolved with the current investigation, were recruited via email and were offered compensation for completing their reviews. The two internal experts were (MJZ and JCV), who had reviewed the outcomes of the qualitative inquiry in Study 1. Each expert independently provided a review of the initial item pool. The goals of expert review were to (1) narrow and refine the initial PIES item pool and (2) attend to content validity.

Procedure

Each expert received an email that provided a description of the research project, our operational definition of emotion invalidation, and instructions for completing the review online via Qualtrics. The instructions and ratings were modeled from Gehlbach and Brinkworth's (2011) guidelines and review form, as well as Lawshe's (1975) work on content validity.

Ratings—Experts rated each item for (1) *relevance*, or how central the item was to the construct of emotion invalidation; (2) *clarity*, or how comprehensible each item was; and (3) *the anticipated mean response* to each item if the survey was administered to a nonclinical sample of college and community participants. Relevance ratings options ranged from 1 to 3, where 1 = *Not necessary*, 2 = *Useful but not essential*, and 3 = *Essential*. Clarity rating options also ranged from 1 to 3 and were 1 = *Not at all understandable*, 2 = *Somewhat understandable*, and 3 = *Extremely understandable*. Experts recorded their expected mean ratings with the same 5-point Likert scale used in the final version of the PIES.

Opportunities to provide qualitative feedback, including an item which asked whether the set of items adequately covered the scope of emotion invalidation as a construct, were available via open-response boxes throughout the survey.

Analytical Approach

Relevance and clarity ratings from each expert were both examined independently and considered in aggregate. Relevance ratings were the primary tool used for deciding whether to retain or exclude an item. Items were retained if they were scored as *Essential* by any

²Scale anchors were inspired by the Difficulties in Emotion Regulation Scale (Gratz & Roemer, 2004).

three experts or by both external reviewers, and excluded if rated *Essential* by only one expert. When item relevance ratings did not fall into the aforementioned categories, we considered redundancy with other items, relevance to participants' narratives in Study 1, and qualitative relevance comments by experts when deciding about retention. Clarity ratings and qualitative clarity comments were examined for items that were retained. If an item had an average clarity score that was less than perfect, wording alterations were considered.

Results and Discussion

Most items received high relevance and clarity ratings upon expert review; thus, 27 of 37 items in the initial PIES item pool were retained in their original or slightly altered form (see Table 2). The experts' expected mean item responses evidenced good variability across individual items, commensurate with our goal to create a measure that could capture a range of variance across both clinical and nonclinical populations. Together, Study 2 provided strong support for the content validity of the PIES item pool, and used expert feedback to finalize the PIES – Preliminary Version (PIES-P) for initial administration. The full results of the expert review are available in our online supplemental material.

Study 3 – Exploration of the PIES-P's Factor Structure

Study 3 was a preliminary psychometric evaluation of the PIES-P, which contained the 27 items that were retained after expert review. We examined convergent validity with two measures of childhood invalidation, internal consistency, and factor structure. A minimum sample size of 300 was selected (Comrey & Lee, 1992; Tabachnick & Fidell, 2006).

Participants and Procedure

A large sample of adults ($N = 402$) completed Study 3 via Qualtrics. Respondents were students ($n = 201$) who participated for course credit and community members recruited and paid 0.75 US dollars through Amazon Mechanical Turk (MTurk; $n = 201$). MTurk workers were eligible if they lived in the US and had a positive record of past work.

Measures

Perceived Invalidation of Emotion Scale – Preliminary Version (PIES-P)—The 27-item PIES-P was assessed as a measure of current emotion invalidation in Study 3. Items were rated on a 5-point Likert scale from 1 (*Almost never; 0–10%*) to 5 (*Almost always; 91–100%*), and averaged together to create a mean invalidation score.

Invalidating Childhood Environment Scale (ICES)—The 18-item ICES (Mountford et al., 2007) is a two-part measure designed to retrospectively assess for emotion invalidation by each parent during childhood. Part 1 of the ICES comprises 14 items about parental behaviors that are rated separately for participants' mothers and fathers. Internal consistency in the Study 3 sample was acceptable ($\alpha = .73$). Part 2 of the ICES was not analyzed in this investigation.

Socialization of Emotion Scale (SES)—The 15-item invalidation subscale of the SES (Sauer & Baer, 2010) was used as a second index of childhood emotion invalidation.

Participants rated items separately for their mother and father using a 7-point Likert scale from 1 (*very unlikely*) to 7 (*very likely*). Internal consistency in the Study 3 sample was excellent ($\alpha = .91$).

Results

Sample Characteristics—The MTurk sample was about half female and married, and mostly White. The student sample was mostly female, White, and unmarried. Student participants were significantly younger, $t(251.31) = 19.58, p < .01$, and more likely to be female, $\chi^2(1) = 13.94, p < .01$.³ See Table 1 for overall sample demographics.

Preliminary Analyses—Most items demonstrated moderate levels of positive skew; however, skewness and kurtosis values were within acceptable ranges for all 27 items in the PIES-P item pool (i.e., skewness < 2 , kurtosis < 4). Correlations between individual items were also examined for the purpose of potentially eliminating items based on redundancy. All items were significantly correlated; however, no items were redundant (i.e., correlation $> .80$) and therefore none were eliminated at this stage.

Exploratory Factor Analysis (EFA)—The factor structure of the PIES-P (see Table 2) was examined using EFA after confirming that the data were appropriate for this technique (Kaiser-Meyer-Olkin = .97, $p < .001$ for Bartlett's Test of Sphericity). Two extraction methods, principal axis factoring and maximum likelihood, were explored. In both cases an oblique rotation was examined because it was expected that any resulting factors would be correlated.

Regardless of extraction method, examination of the results of our Scree test (Cattell, 1966) and a parallel analysis using the Monte Carlo PCA for Parallel Analysis program (Watkins, 2006) indicated that the PIES-P was unidimensional. Factor 1 explained 57.03% of the variance. Examination of the results using the Kaiser-Guttman criterion suggested the presence of two additional factors, which explained an additional 4.82% and 3.93% of the variance. Factor 2, which was the same across principal axis factoring and maximum likelihood, consisted of 3 items (17,18,19) that used the same item stem. Factor 3 had two items (PAF: 6,7, or ML: 25, 26). Given that the PIES-P was best characterized as unidimensional, we report factor weights with all 27 PIES-P items constrained to a single factor in Table 2.

Internal Consistency—The 27 PIES-P items were averaged to create a composite score of current invalidation, which evidenced excellent internal consistency (Cronbach's $\alpha = .97$).

Convergent validity—The PIES-P ($M = 1.90, SD = 0.77$) positively correlated with both the ICES ($M = 2.57, SD = 0.40; r = .43, p < .01$) and the SES ($M = 3.08, SD = 1.03; r = .35, p < .01$).³ The strength of the correlations did not suggest redundancy, and the two childhood

³Independent ratings of mothers and fathers on the ICES and SES were averaged to create a composite score for each measure prior to analysis. If either parent was reportedly uninvolved during childhood, only the average score for the involved parent was used in analysis.

invalidation measures correlated more strongly with one another ($r = .57, p < .01$). There were no sample differences or sex differences in mean PIES-P scores (all $ps > .05$).

Discussion

Study 3 suggested that the 27-item PIES-P was a unidimensional measure that demonstrated good convergent validity with existing measures of childhood invalidation while remaining distinct. Study 3 therefore provided emerging evidence as to the psychometric properties of the PIES-P, which were expanded upon in Study 4.

Study 4 - Confirmatory Factor Analysis and Finalization of the PIES

Study 4 aimed to confirm the unidimensional factor structure of the PIES-P through confirmatory factor analysis (CFA) of the measure structure in a new participant sample. We selected a minimum sample size of 600 participants to create a primary dataset ($n = \sim 300$) on which to run our CFA and a secondary reserve dataset ($n = \sim 300$) for a second CFA if changes to the measure were needed after analysis of the primary dataset. As with Study 3, all participants ($N = 604$) were either university students ($n = 301$) compensated with course credit or MTurk workers ($n = 303$) paid 0.25 US dollars to complete measures online using Qualtrics.

Results

Sample Characteristics—Similar to Study 3, MTurk workers were about half female and married, and were primarily White, while students were primarily female, White, and unmarried. Students were significantly younger, $t(338.66) = 23.04, p < .001$ and more likely to be White, $\chi^2(1) = 4.41, p = .04$, than MTurk workers. See Table 1 for overall sample demographics.

Confirmatory Factor Analysis (CFA)—Cases were randomly split between two datasets prior to analyses. Dataset 1 comprised 295 participants and Dataset 2 comprised 309 participants (160 and 143 from MTurk respectively). The unidimensional factor structure of the PIES-P found in Study 3 was imposed on participants' responses in Dataset 1 using CFA. Evaluation of fit indices indicated an unacceptable model fit: $\chi^2(324) = 1178.69, p < .001$, CMIN/DF = 3.64, GFI = .77, NFI = .81, CFI = .86, RMSEA = .095 (CI: .09, .10), and AIC = 1286.69.⁵ Model modification indices and item regression weights were thus examined. Two additional models, with minor changes based on these indices and on the Study 3 findings that similarly worded items might cluster together, were examined, but model fit was still poor. Given that larger models may be statistically more difficult to fit, we re-examined the PIES-P item pool with the goal of reducing the total number of items by eliminating redundancy. This was accomplished by considering (1) factor loadings from Study 3, (2) inter-item correlations, (3) conceptual redundancy with other items, and (4) preserving representation of elements of the prominent thematic codes identified in Study 1. Items judged to be strong contributors to the measure, both statistically and theoretically,

⁵We expected good model fit to be represented by a nonsignificant χ^2 goodness-of-fit test, CMIN/DF < 2, GFI .95, NFI .95, CFI .95, and RMSEA upper confidence interval value .08 (Tabachnick & Fidell, 2006). For the purposes of model comparison, the Akaike Information Criterion (AIC) values were also examined, with values closer to zero being more favorable.

were retained. The inter-item correlations between retained items were all less than $|r| = 0.70$.

After revisions, the finalized Perceived Invalidation of Emotion Scale (PIES) consisted of 10 items. For the purposes of CFA, these were split into two second-order factors reflecting question stem, which were thought to contribute to one higher-order factor and were thus modeled as correlated. One second-order factor consisted of items 2, 9–11, 14, and 16 (i.e., items beginning with “When I share how I’m feeling...”). The second second-order factor consisted of items 22–24 and 26 (i.e., items beginning with “Others...”). The fit indices for this model were substantially improved from the initial model and were consistent with a well-fitting model. Specifically, $\chi^2(34) = 52.37$, $p = .02$, CMIN/DF = 1.54, GFI = .97, NFI = .97, CFI = .99, RMSEA = .043 (CI: .016, .065), and AIC = 94.37. Having achieved good model fit in Dataset 1, the aforementioned model was examined using the independent sample ($n = 309$) in Dataset 2. The fit indices for the final model in the independent sample confirmed that model fit was good: $\chi^2(34) = 567.34$, $p < .01$, CMIN/DF = 1.98, GFI = .96, NFI = .97, CFI = .98, RMSEA = .056 (CI: .036, .076), and AIC = 109.34. The internal consistency of the finalized 10-item PIES was excellent in both samples ($\alpha = .94$ for Dataset 1 and $\alpha = .93$ for Dataset 2). The finalized measure is available in Appendix A.

Discussion

Substantial changes to the 27-item PIES-P were made after attempts to confirm its factor structure revealed problematic fit indices, despite a strong item pool evidencing high factor loadings. After several unsuccessful attempts to improve model fit via minor revisions, more major changes were necessary; the total number of items was reduced by more than half and two lower order, correlated factors (based on item wording) were modeled. We achieved excellent model fit, confirmed in an independent sample. The revisions resulted in the finalized Perceived Invalidation of Emotion Scale (PIES), a statistically strong measure that is brief enough to be administered in just a few minutes, and thus is practical for both research and clinical purposes.

Study 5 - Psychometric Characteristics of the Finalized PIES

The purpose of Study 5 was to examine the psychometric properties of the 10-item PIES. A short-term longitudinal design with a one-month follow-up period was used to examine the internal consistency, test-retest reliability, and validity (convergent, divergent, incremental, and predictive) of responses on the PIES.

Participants and Procedure

As in previous studies, participants were either US adults recruited through MTurk ($n = 100$) or university students ($n = 99$). All participants completed baseline measures (Time 1) and were sent instructions via email for completing Time 2 measures ~30 days later. They had 8 days to submit their responses. MTurk workers were paid 2 US dollars and students were provided with research credit at each time point completed.⁶

A total of 175 participants provided usable data at both time points (retention rates 94.9% for students and 81% for MTurk). The average time between Time 1 and Time 2

participation was 33.07 days (range = 27.06–38.16). The follow-up period for student participants ($M = 33.50$ days, $SD = 2.37$) was on average one day longer than for MTurk participants ($M = 32.58$ days, $SD = 1.65$), a difference that was statistically significant, $t(165.91) = 2.99, p < .01$.⁴

Measures

Current Invalidation—The 10-item PIES was used to assess current (past month) emotion invalidation at both Time 1 and Time 2. Respondents rate items on a 5-point Likert scale as previously described. The reading level of the measure is appropriate for use with general adult samples (6.6 Flesch-Kincaid Grade Level for item pool and instructions per Microsoft Word). Responses were averaged to create a separate mean emotion invalidation score for each time point. Internal consistency at both time points was excellent ($\alpha_{T1} = .91, \alpha_{T2} = .93$).

Childhood Invalidation—The Invalidating Childhood Environment Scale (ICES; Mountford et al., 2007) and the Socialization of Emotion Scale (SES; Krause et al., 2003) were administered at Time 1 to assess recollections of childhood invalidation (as in Study 3). Internal consistency was questionable/poor for the ICES ($\alpha = .60$) and good for the SES ($\alpha = .88$).

Emotional Functioning—Two measures were administered at Times 1 and 2 to assess emotional functioning. The Depression Anxiety Stress Scales (DASS-21; J. D. Henry & Crawford, 2005) composite score is a measure of general emotional distress which is strongly and positively correlated with measures of negative affect. Respondents rate items on a 4-point Likert scale from 0 (*did not apply to me at all*) to 3 (*applied to me very much or most of the time*), with higher summed scores indicating greater distress. Internal consistency in the current sample was excellent (Time 1 $\alpha = .92$, Time 2 $\alpha = .93$).

The 36-item Difficulties with Emotion Regulation Scale (DERS; Gratz & Roemer, 2004) indexes emotion dysregulation. Respondents rate the applicability of each item to their experiences on a 5-point Likert scale from 1 (*almost never; 0–10%*) to 5 (*almost always; 91–100%*), with greater scores indicating more difficulty regulating emotion. Internal consistency was excellent (Time 1 $\alpha = .93$; Time 2 $\alpha = .94$).

The 10-item McLean Screening Instrument for Borderline Personality Disorder (MSI-BPD; Zanarini et al., 2003) was completed at Time 1 only and uses true/false items to assess for the presence of borderline symptoms. True items are summed to create a total score where higher scores are indicative of more symptoms. Internal consistency was good ($\alpha = .80$).

Social Functioning—Two measures were administered at Time 1 to assess social functioning. The 6-item Social Support Questionnaire - Short Form (SSQ6; Crowne & Marlowe, 1960; Sarason, Sarason, Shearin, & Pierce, 1987) is a scenario-based measure that

⁶Time 1 data from 7 of the original MTurk workers was indicative of random responding. These workers were not compensated, per policy for low-quality work. These 7 worker assignments were thereafter made available to new workers to maintain the total desired sample of 100 community participants.

⁴Adjusted t -test values are reported due to lack of equal variance between groups.

is used to index both number of social supports (values from 0–9 per scenario) and social support satisfaction (rated on a Likert scale from 1 (*very dissatisfied*) to 6 (*very satisfied*) per scenario). Responses are averaged across all of the scenarios. Internal consistency was excellent for both number of supports and support satisfaction ($\alpha = .94$ and $\alpha = .95$ respectively).

The 10-item Marlowe-Crowne Social Desirability Scale - Short Form, Version XI (MCSF; Crowne & Marlowe, 1960; Reynolds, 1982) indexes individuals' tendencies to present themselves in a positive light, and was included to assess social desirability as a response tendency which could confound our results. Items are presented in a *true* or *false* response format and describe culturally approved or disapproved behaviors which for which the socially desirable response in actuality has a low incidence of occurrence.

Health—The 26-item World Health Organization Quality of Life - Brief (WHOQOL; The WHOQOL Group, 1998) measures individuals' perceived quality of life in the physical, social, psychological, and environmental health domains. Items are rated on a Likert scale from 1 to 5, with verbal scale anchors that change depending upon the item content; higher scores indicate greater health. The WHOQOL was administered at both Times 1 and 2. Internal consistency was adequate or good for all subscales at both time points (α s = .73–.85).

Personality—The Big Five Inventory (BFI; John, Donohue, & Kentle, 1991) is widely used to measure the five primary personality domains. Participants rate items on a 5-point Likert scale from 1 (*disagree strongly*) to 5 (*agree strongly*). It was given at Time 1 to index neuroticism, agreeableness, and openness. Internal consistency was good for the neuroticism subscale ($\alpha = .89$) and adequate for the agreeableness ($\alpha = .79$) and openness subscales ($\alpha = .77$).

Data Preparation

Data quality for the full sample was examined prior to analysis. Fourteen participants were excluded because of low effort (i.e., random responding, very short response times coupled with limited response variability), leaving a final sample of 161 participants (87 students) for use in analyses. All variables were within acceptable limits for skewness and kurtosis, and data did not violate assumptions of multicollinearity.

Results

Sample Characteristics—Study 5 participants were similar to those in Studies 3 and 4 (Table 1). MTurk workers were about half female, and primarily White, and unmarried. Student participants were primarily female, White, and unmarried. As before, student participants were significantly younger than MTurk workers, $t(76.91) = 13.48, p < .001^4$, and were significantly more likely to be female, $\chi^2(1) = 5.34, p = .02$.

Student participants and MTurk workers were largely similar on study variables. However, student participants reported greater childhood emotion invalidation on the ICES ($p = .04$) and evidenced greater neuroticism ($p < .01$) and problems with emotion regulation ($p < .01$)

at Times 1 and 2). Yet, students also reported more social supports ($p < .01$) and greater environmental health ($p < .01$ at Times 1 and 2). There were no significant sex differences aside from women reporting greater neuroticism ($p < .01$).

Psychometric Properties of the PIES

Test-Retest Reliability—A Pearson correlation coefficient was used to examine test-retest reliability. Responses on the PIES demonstrated good test-retest reliability, as evidenced by a moderately large correlation ($r = .67, p < .01$) between Times 1 and 2 scores.

Convergent Validity—We examined associations between the PIES and the two childhood invalidation measures to preliminarily assess convergent validity. Though the correlations were somewhat smaller than expected ($r = .18, p = .02$ for ICES; $r = .27, p < .01$ for SES) they were statistically significant, suggesting modest convergence. We also examined associations between the PIES and current personality factors (neuroticism, agreeableness) and social support satisfaction, which were expected to evidence small, but significant, positive or negative associations with responses on the PIES. Results revealed a significant positive correlation between Time 1 PIES scores and neuroticism ($r = .34, p < .01$). There were also significant negative correlations between Time 1 PIES scores and agreeableness ($r = -.37, p < .01$) and social support satisfaction ($r = -.38, p < .01$).

Divergent Validity—PIES scores were compared with scores on measures of constructs that were expected to show small or negative correlations with emotion invalidation. As anticipated, correlations between Time 1 PIES scores and openness ($r = -.02, p = .79$) and number of social supports ($r = -.07, p = .38$) were not significant.

Concurrent Validity—Concurrent validity was assessed by examining correlations between the PIES and measures associated with both psychopathology and health when measured at the same time point. As expected, greater emotion invalidation was significantly correlated with higher levels of all variables related to psychopathology (emotional distress and dysregulation) and lower levels of all variables related to health (all $ps < .01$; Table 3).

Incremental Validity—Hierarchical regression was used to examine whether emotion invalidation as indexed by the PIES would predict outcomes beyond what can be accounted for by general childhood invalidation. Time 1 scores were used for all analyses. Three separate hierarchical regression analyses examined emotion dysregulation (DERS total scores), borderline features (MSI-BPD total scores), and emotional distress (DASS-21 total scores) as outcomes. Predictor variables were the same in all three models; childhood invalidation as indexed by the SES, sample type (student = 0, MTurk = 1), and social desirability (MCSF total scores) were entered in Step 1. Only one childhood invalidation measure was entered in Step 1 because of the strong correlation between the two measures ($r = .53, p < .01$); the SES was selected because it had the greatest zero-order correlation with the PIES. Sample was included as a Step 1 variable to control for the sample differences evidenced in preliminary analyses. The PIES was then entered at Step 2.

The three regression analyses are presented in Table 4. Social desirability and childhood invalidation significantly predicted all outcomes at Step 1; however, the PIES accounted for

an additional 25% of the variance in current distress, 12% of the variance in emotion dysregulation, and 8% of the variance in borderline features when added to the model, providing preliminary evidence of incremental validity. However, because recent simulation studies indicate that Type I error rates can be substantially inflated when using observed variables in incremental validity analyses (Westfall & Yarkoni, 2016), these analyses should be viewed tentatively.

Predictive Validity—We examined correlations between Time 1 PIES scores and emotional functioning and health status at Time 2 as a preliminary investigation of the measure's predictive validity. As expected, PIES scores at Time 1 were positively associated with emotional distress on the DASS-21 and emotion dysregulation on the DERS at Time 2 ($r = .51$ and $r = .40$ respectively, both $ps < .01$), as well as negatively associated with all health status variables ($r_{\text{physical}} = -.37$, $r_{\text{environmental}} = -.43$, $r_{\text{relational}} = -.49$, $r_{\text{psychological}} = -.51$, all $ps < .01$).

Hierarchical regression analyses were then used to examine the ability of the PIES to predict change in symptom measures over time (Table 5). Current distress (DASS-21 total scores), relational health, and psychological health (WHOQOL subscale scores) at Time 2 were examined as outcomes. In each regression model, the Time 1 scores for the commensurate measure were entered in Step 1 to control for existing symptoms. PIES Time 1 scores were entered in Step 2. Results of these final analyses were mixed. The PIES did not predict emotional distress at Time 2 when emotional distress was entered in the model previously at Time 1 ($p = .10$). However, greater currently invalidation on the PIES predicted reduced relational health ($p < .01$) and reduced psychological health ($p < .01$) when Time 1 scores on the requisite subscales were included.

Discussion

Study 5 was the culminating study in this series of investigations, which endeavored to design and examine the psychometric properties of a new measure of perceived emotion invalidation. This study provides preliminary psychometric support for the PIES. The internal consistency of the measure was excellent. Test-retest reliability was high, but the moderately large correlation between Times 1 and 2 scores also suggests that PIES scores change over time and therefore that the measure is indeed more of a state, rather than trait, measure. The PIES also converged with measures of theoretically related constructs in the expected direction (i.e., evidenced small-to-moderate positive associations with childhood invalidation and neuroticism and small-to-moderate negative associations with agreeableness and social support satisfaction) and diverged from measures that were not expected to be associated with emotion invalidation (i.e., number of social supports and openness). Correlations with existing measures of childhood invalidation were weaker than anticipated; however, the PIES intentionally deviated from a focus on past experiences and to a focus on current emotion invalidation. The results of Study 5 confirm that emotion invalidation is associated with psychopathology/emotional distress and health status. While the PIES was only able to predict relatively small increases in additional symptomology at Time 2 when controlling for Time 1 symptomology, the strength of the relations between Time 1 and Time 2 scores of the same measures were so high that predicting an additional 2–3% variance may

be clinically meaningful. Present symptoms are clearly a very strong predictor of future symptoms, emotionally invalidating experiences may add additional fuel to the fire, so to speak, for individuals already in emotional distress.

General Discussion

The lack of validated instruments for measuring current emotion invalidation has prevented in-depth investigation of emotion invalidation's role in the development of psychopathology and the exacerbation of health problems. This series of studies resulted in the creation of the Perceived Invalidation of Emotion Scale (PIES), a promising and practical tool to accelerate research on this previously elusive construct.

Current Perceived Emotion Invalidation

We based the PIES on a revised and clarified definition of emotion invalidation. Our definition focused on *emotion* invalidation, beyond other types of invalidation (e.g., criticism; Cheavens et al., 2005). We focused on emotions because emotions are central to the human experience, and because emotion invalidation is central to conceptualizations of clinical phenomena, most notably borderline personality disorder (Linehan, 1993). The definition we advanced also emphasized *perception*, rather than observable behavior, was an intentional choice that has both advantages and disadvantages. On the one hand, perceptions may be inaccurate; someone may *feel* invalidated even if the other person had good intentions or provided what would appear to be support to an outside observer. Perceptions of the world are often skewed, particularly for those with clinical conditions who may inadvertently attend to negative information (Baer, Peters, Eisenlohr-Moul, Geiger, & Sauer, 2012; Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Winer & Salem, 2016). Conversely, regardless of the objective nature of the person deemed to be invalidating, if an individual perceives invalidation, that might be sufficient to influence outcomes. Self-report symptom measures, which prioritize individuals' experiences of symptoms, are commonly administered in self-report format, even in rigorous treatment trials (e.g., Cuijpers, Straten, & Warmerdam, 2007). Focusing on perceptions is consistent with rich literatures emphasizing the importance of social-cognitive variables such as self-efficacy, expectancies and beliefs (Bandura, 2001) in predicting behavior. Future research exploring the relation between observable invalidating behavior and perceptions of emotion invalidation is certainly warranted. However, we would hypothesize that perceptions (felt experiences) of social responses to shared emotions may be even more important in predicting negative outcomes than the observable behavior itself.

Results presented here align with previous research which has found that invalidation seems to associate with a variety of negative health indicators (i.e., physical, psychological, social, and environmental) and clinical constructs (e.g., borderline features, emotion dysregulation, and emotional distress). Moreover, we found evidence that emotion invalidation as indexed by the PIES predicted worsened relational and psychological health over a one-month follow-up period, even when controlling for baseline health. Together, our results provide evidence for the utility of the PIES as a measure of perceived emotion invalidation.

Strengths and Limitations

Development of the PIES followed a sequential, theoretically grounded process which had many strengths. We developed our items based on individuals' experiences with emotion invalidation, as explored via a qualitative study. Participant responses in Study 1 suggested that a wide variety of reactions to emotion can be experienced as invalidating. By coding participants' narrative descriptions of their experiences, we were able to underscore elements of invalidation that align with past research and theory (e.g., responses captured by the direct invalidation code, such as "Don't be upset, you have no reason to be upset."). Even more importantly, we revealed elements of invalidation that would not have been included based on prior theory, such as the perception that others' not mirroring or matching shared emotions was experienced as invalidating. This approach, along with emphasizing participants' wording when developing items, resulted in a measure that is grounded in the experiences of real people who have felt their emotional responses were deemed as inaccurate or inappropriate by others. Although there were ultimately fewer PIES items than qualitative codes, we believe that the experiences that participants described within each of the qualitative codes—even when not explicitly represented (e.g., the invalidation by group membership code)—were likely captured within the 10 items ultimately included in the PIES. With the addition of having items vetted by expert reviewers, the development of the PIES was thus grounded in both individual experience and in psychological knowledge.

Additional strengths of the current investigation include the purposeful recruitment of both college student and community samples for each study. Extant research on invalidation has primarily relied on either college student participants (Robertson et al., 2013; Sauer & Baer, 2010; Shenk & Fruzzetti, 2011; Woodberry & Popenoe, 2008) or specific clinical populations (e.g., Mountford et al., 2007; Sells et al., 2008). The present investigation extends beyond this previous work by aiming to design a measure of perceived emotion invalidation that is more broadly applicable to adults of various ages and educational backgrounds. Future research would benefit from examining the clinical utility of the PIES, including in specific patient samples (e.g., individuals with a diagnosis of BPD) or among individuals with varying clinical vulnerabilities (e.g., people with high emotional reactivity, people with emotion regulation difficulties). Additional work regarding generalizability is also needed due to the limited data on individual difference characteristics collected in this study and the lack of racial diversity in our sample (as described in our discussion of limitations below).

Importantly, the current study followed traditional psychometric processes for scale design, including both EFA and a two-step CFA to verify altered scale structure. We ultimately conclude that emotion invalidation is a broad, multifaceted social experience and one that is best measured as unidimensional; despite the variety of discrete responses to emotion included within the PIES, no meaningful subscales emerged in the exploratory and confirmatory factor analyses conducted in Studies 3 and 4. This suggests that the measures of related, but narrow, constructs (e.g., criticism; Cheavens et al., 2005), which have been used in past research, likely do not capture the full scope of emotion invalidation and should not be used in isolation when measuring invalidation in future research.

The results of this investigation should be interpreted in light of relevant limitations which include the use of self-report data and reliance on convenience samples. Diversity, more generally, was also limited among participants in this investigation. This is problematic given that perceptions of emotions may be culturally based (Gendron, Roberson, van der Vyver, & Barrett, 2014; Tamir et al., 2015), and norms and sensitivities to responses to emotions may differ across cultures and subgroups. Future studies should also expand the sample diversity and consider whether individual difference factors (e.g., race, ethnicity, cultural background, sexual orientation) influence experiences of emotion invalidation.

The structure and wording of the PIES is also relevant to its potential applications. In current form, the instructions for the PIES asks participants to consider *how often* others with whom they have regular contact respond to their shared emotions in the manner described. We used this approach because we endeavored index individuals aggregate levels of emotion invalidation and therefore treated emotion invalidation as a general environmental characteristic rather than nesting it within any particular relationship(s). Modifications to the instructions may need to be made (and assessed for impact) based on the intended use (e.g., if a particular research question required assessment of emotion invalidation by a romantic partner specifically).

Conclusion and Future Directions

Possibly the greatest contribution of this investigation is the potential for future research on a construct that has garnered relatively little attention despite theorized importance. Beyond those areas already discussed, future research could include examining the relation between emotion invalidation and self-invalidation, as the biosocial theory of BPD hypothesizes that experiencing emotion invalidation ultimately leads individuals to begin to invalidate themselves (Linehan, 1993). It is possible that these processes could be mechanisms by which emotion invalidation influences outcomes. Future research could also examine emotion invalidation as a potential intervention point for psychological and physical health practices; for example, such work could examine whether decreasing perceived emotion invalidation through either modification of perceptions (e.g., via promotion of acceptance-oriented strategies) or modification of the individuals' social environments (e.g., via family interventions or working to end unhealthy relationships) has any effect on health. Our articulation of a definition of emotional invalidation and development of the PIES together provides a base from which to launch this seemingly promising area of study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: Execution of this work was supported by a grant from the Marie Wilson Howells Foundation, awarded to the first author. Manuscript preparation was supported by grants T32DA022981 (PI: Kilts) and R25DA037190 (PI: Beckwith) which support the first author.

We would like to thank Dr. Ana Bridges and Dr. Denise Beike, who provided feedback on elements of this project, Mary Garaudy and Elise Warner for assistance with qualitative coding, and Mary Cantrell and DeAnn Hubberd for

their feedback on written elements of this manuscript. Of note, we would also like to extend our sincere thanks to Dr. Chad Shenk and Dr. Bonnie Sturrock for providing their expertise in Study 2 expert reviews.

References

- Baer RA, Peters JR, Eisenlohr-Moul TA, Geiger PJ, Sauer SE. Emotion-related cognitive processes in borderline personality disorder: A review of the empirical literature. *Clinical Psychology Review*. 2012; 32:359–369. DOI: 10.1016/j.cpr.2012.03.002 [PubMed: 22561966]
- Bandura A. Social cognitive theory: An agentic perspective. *Annual Review of Psychology*. 2001; 52:1–26. DOI: 10.1146/annurev.psych.52.1.1
- Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, van IJzendoorn MH. Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*. 2007; 133:1–24. DOI: 10.1037/0033-2909.133.1.1 [PubMed: 17201568]
- Cano A, Leong LEM, Williams AM, May DKK, Lutz JR. Correlates and consequences of the disclosure of pain-related distress to one's spouse. *Pain*. 2012; 153:2441–7. DOI: 10.1016/j.pain.2012.08.015 [PubMed: 23059054]
- Cattell RB. The scree test for the number of factors. *Multivariate Behavioral Research*. 1966; 1:245–276. DOI: 10.1207/s15327906mbr0102_10 [PubMed: 26828106]
- Center for Behavioral Health Statistics and Quality. Behavioral health trends in the United States: Results from the 2014 National Survey on Drug Use and Health. 2015. (HHS Publication No. SMA 15-4927, NSDUH Series H-50). Retrieved from <http://www.samhsa.gov/data/>
- Cheavens JS, Zachary Rosenthal M, Daughters SB, Nowak J, Kosson D, Lynch TR, Lejuez CW. An analogue investigation of the relationships among perceived parental criticism, negative affect, and borderline personality disorder features: the role of thought suppression. *Behaviour Research and Therapy*. 2005; 43:257–68. DOI: 10.1016/j.brat.2004.01.006 [PubMed: 15629754]
- Clark L, Watson D. Constructing validity: Basic issues in objective scale development. *Psychological Assessment*. 1995; 7:309–319.
- Crowell SE, Beauchaine TP, Linehan MM. A biosocial developmental model of borderline personality: Elaborating and extending Linehan's theory. *Psychological Bulletin*. 2009; 135:495–510. DOI: 10.1037/a0015616 [PubMed: 19379027]
- Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*. 1960; 24:349–354. DOI: 10.1037/h0047358 [PubMed: 13813058]
- Comrey AL, Lee HB. *A first course in factor analysis*. Hillsdale NJ: Erlbaum; 1992.
- Cuijpers P, Van Straten A, Warmerdam L. Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review*. 2007; 27:318–326. DOI: 10.1016/j.cpr.2006.11.001 [PubMed: 17184887]
- Ekman P, Friesen W. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*. 1971; 17:124–129. [PubMed: 5542557]
- Elfenbein H, Ambady N. On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*. 2002; 128:203–235. DOI: 10.1037//0033-2909.128.2.203 [PubMed: 11931516]
- Ford G, Waller G, Mountford V. Invalidating childhood environments and core beliefs in women with eating disorders. *European Eating Disorders Review*. 2011; 19:316–21. DOI: 10.1002/erv.1053 [PubMed: 20957770]
- Gehlbach H, Brinkworth ME. Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology*. 2011; 15:380–387. DOI: 10.1037/a0025704
- Gendron M, Roberson D, van der Vyver JM, Barrett LF. Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion*. 2014; 14:251–62. DOI: 10.1037/a0036052 [PubMed: 24708506]
- Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*. 2004; 26:41–54. DOI: 10.1023/B:JOBA.0000007455.08539.94

- Haslam M, Arcelus J, Farrow C, Meyer C. Attitudes towards emotional expression mediate the relationship between childhood invalidation and adult eating concern. *European Eating Disorders Review*. 2012; 20:510–514. DOI: 10.1002/erv.2198 [PubMed: 22933402]
- Haslam M, Mountford V, Meyer C, Waller G. Invalidating childhood environments in anorexia and bulimia nervosa. *Eating Behaviors*. 2008; doi: 10.1016/j.eatbeh.2007.10.005
- Henry JD, Crawford JR. The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology*. 2005; 44:227–39. DOI: 10.1348/014466505x29657 [PubMed: 16004657]
- John OP, Donahue EM, Kentle RL. *The Big Five Inventory – Versions 4a and 54*. Berkeley, CA: University of California at Berkeley, Institute of Personality and Social Research; 1991.
- Kool MB, Geenen R. Loneliness in patients with rheumatic diseases: The significance of invalidation and lack of social support. *The Journal of Psychology*. 2012; 146:229–241. [PubMed: 22303622]
- Kool MB, Van Middendorp H, Lumley MA, Bijlsma JW, Geenen R. Social support and invalidation by others contribute uniquely to the understanding of physical and mental health of patients with rheumatic diseases. *Journal of Health Psychology*. 2013; 18:86–95. DOI: 10.1177/1359105312436438 [PubMed: 22363049]
- Kool MB, van Middendorp H, Boeije H, Geenen R. Understanding the lack of understanding: Invalidation from the perspective of the patient with fibromyalgia. *Arthritis & Rheumatism*. 2009; 61:1650–1656. DOI: 10.1002/art.24922 [PubMed: 19950317]
- Kool MB, van Middendorp H, Lumley MA, Schenk Y, Jacobs JW, Bijlsma JW, Geenen R. Lack of understanding in fibromyalgia and rheumatoid arthritis: The Illness Invalidation Inventory (3*I). *Annals of the Rheumatic Diseases*. 2010; 69:1990–1995. DOI: 10.1136/ard.2009.123224 [PubMed: 20498203]
- Krause ED, Mendelson T, Lynch TR. Childhood emotional invalidation and adult psychological distress: The mediating role of emotional inhibition. *Child Abuse & Neglect*. 2003; 27:199–213. DOI: 10.1016/S0145-2134(02)00536-7 [PubMed: 12615094]
- Lawshe CH. A quantitative approach to content validity. *Personnel Psychology*. 1975; 28:563–575.
- Leong LE, Cano A, Johansen AB. Sequential and base rate analysis of emotional validation and invalidation in chronic pain couples: Patient gender matters. *The Journal of Pain*. 2011; 12:1140–1148. DOI: 10.1016/j.pain.2011.04.004 [PubMed: 21703939]
- Linehan MM. *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford Press; 1993.
- Linehan MM. Self-verification and drug abusers: Implications for treatment. *Psychological Science*. 1997; 8:181–183.
- Linton SJ, Boersma K, Vangronsveld K, Fruzzetti A. Painfully reassuring? The effects of validation on emotions and adherence in a pain test. *European Journal of Pain*. 2012; 16:592–599. DOI: 10.1016/j.ejpain.2011.07.011 [PubMed: 22396087]
- Mountford V, Corstorphine E, Tomlinson S, Waller G. Development of a measure to assess invalidating childhood environments in the eating disorders. *Eating Behaviors*. 2007; 8:48–58. DOI: 10.1016/j.eatbeh.2006.01.003 [PubMed: 17174851]
- Reeves M, James LM, Pizzarello SM, Taylor JE. Support for Linehan’s biosocial theory from a nonclinical sample. *Journal of Personality Disorders*. 2010; 24:312–326. DOI: 10.1521/pedi.2010.24.3.312 [PubMed: 20545497]
- Reise SP, Waller NG, Comrey AL. Factor analysis and scale revision. *Psychological Assessment*. 2000; 12:287–297. DOI: 10.1037//1040-3590.12.3.287 [PubMed: 11021152]
- Reynolds WM. Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*. 1982; 38:119–125.
- Robertson CD, Kimbrel Na, Nelson-Gray RO. The Invalidating Childhood Environment Scale (ICES): Psychometric properties and relationship to borderline personality symptomatology. *Journal of Personality Disorders*. 2013; 27:402–410. DOI: 10.1521/pedi_2012_26_062 [PubMed: 22984861]
- Saldaña J. *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage; 2013.
- Sarason IG, Sarason BR, Shearin EN, Pierce GR. A brief measure of social support: Practical and theoretical implications. *Journal of Social and Personal Relationships*. 1987; 4:497–510. DOI: 10.1177/0265407587044007

- Sauer SE, Baer RA. Validation of measures of biosocial precursors to borderline personality disorder: Childhood emotional vulnerability and environmental invalidation. *Assessment*. 2010; 17:454–466. DOI: 10.1177/1073191110373226 [PubMed: 20644079]
- Sells D, Black R, Davidson L, Rowe M. Beyond generic support: Incidence and impact of invalidation in peer services for clients with severe mental illness. *Psychiatric Services*. 2008; 59:1322–1327. DOI: 10.1176/appi.ps.59.11.1322 [PubMed: 18971409]
- Shariff A, Tracy J. What are emotion expressions for? *Current Directions in Psychological Science*. 2011; 20:395–399. DOI: 10.1177/09637214111424739
- Shenk CE, Fruzzetti AE. The impact of validating and invalidating responses on emotional reactivity. *Journal of Social and Clinical Psychology*. 2011; 30:163–183. DOI: 10.1521/jscp.2011.30.2.163
- Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 5. Needham Heights, MA: Allyn & Bacon; 2006.
- Tamir M, Schwartz SH, Cieciuch J, Riediger M, Torres C, Scollon C, Vishkin A. Desired emotions across cultures: A value-based account. *Journal of Personality and Social Psychology*. 2015; :1–16. DOI: 10.1037/pspp0000072
- The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine*. 1998; 28:551–558. DOI: 10.1017/S0033291798006667 [PubMed: 9626712]
- Vangronsveld KL, Linton SJ. The effect of validating and invalidating communication on satisfaction, pain and affect in nurses suffering from low back pain during a semi-structured interview. *European Journal of Pain*. 2012; 16:239–346. DOI: 10.1016/j.ejpain.2011.07.009 [PubMed: 22323376]
- Watkins MW. Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods*. 2006; 5:344–346.
- Westfall J, Yarkoni T. Statistically controlling for confounding constructs is harder than you think. *PLoS ONE*. 2016; 11(3):e0152719. [PubMed: 27031707]
- Winer ES, Salem T. Reward devaluation: Dot-probe meta-analytic evidence of avoidance of positive information in depressed persons. *Psychological Bulletin*. 2016; 142:18–78. DOI: 10.1037/bul0000022 [PubMed: 26619211]
- Woodberry KaPopenoe EJ. Implementing Dialectical Behavior Therapy with adolescents and their families in a community outpatient clinic. *Cognitive and Behavioral Practice*. 2008; 15:277–286. DOI: 10.1016/j.cbpra.2007.08.004
- Woodberry KaGallo KP, Nock MK. An experimental pilot study of response to invalidation in young women with features of borderline personality disorder. *Psychiatry Research*. 2008; 157:169–80. DOI: 10.1016/j.psychres.2007.06.007 [PubMed: 17913240]
- Zanarini MC, Vujanovic AA, Parachini EA, Boulanger JL, Frankenburg FR, Hennen J. A screening measure for BPD: The McLean Screening Instrument for Borderline Personality Disorder (MSI-BPD). *Journal of Personality Disorders*. 2003; 17:568–573. [PubMed: 14744082]

Appendix A: The Perceived Invalidation of Emotion Scale

Instructions

Please take a moment to think about your relationships with the people who you are in contact with on a regular basis (i.e., at least once per week) and how they respond to **your emotions** when you share them. You may want to consider your relationships with family, friends, intimate partners, coworkers, and acquaintances.

Then, please indicate how often each item applied to you over the past month using the following scale:

1	2	3	4	5
Almost Never (0–10%)	Sometimes (11–35%)	About half the time (36–65%)	Most of the time (66–90%)	Almost Always (91–100%)
____ 1. When I share how I'm feeling, others don't seem to mirror or match my emotions. For example, they don't share sadness with me when I'm sad or happiness with me when I'm happy.				
____ 2. When I share how I'm feeling, others want me to "get over it" or "accept it and move on."				
____ 3. When I share how I'm feeling, others seem like they don't want to hear what I have to say.				
____ 4. When I share how I'm feeling, others look down on me or judge me.				
____ 5. When I share how I'm feeling, others don't take me seriously.				
____ 6. When I try to share how I'm feeling, others tell me or imply what I should actually feel.				
____ 7. Others get mad or upset at me when I express my feelings.				
____ 8. Others don't take my side or agree with how I'm feeling.				
____ 9. Others make me feel like it's not okay for me to feel the way that I do.				
____ 10. Others make me feel that my emotions are unimportant.				

Public Significance Statement

This series of five studies resulted in the creation and validation of the first measure of current emotion invalidation (the PIES). Emotion invalidation as assessed by the PIES was linked to worse emotional functioning and lower quality of life. The measure will help researchers learn more about how perceptions of social experiences affect health.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Demographic Data, Separately by Study.

	Study 1 <i>N</i> = 22	Study 3 <i>N</i> = 402	Study 4 <i>N</i> = 604	Study 5 <i>N</i> = 161
<i>Characteristics</i>	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)
Age (<i>Mean</i> (<i>SD</i>))	28.86 (13.93) ^a	29.19 (13.41) ^b	27.98 (12.40) ^a	25.89 (9.43) ^b
Sex				
Female	13 (59.1%)	256 (63.7%)	358 (59.3%)	94 (58.4%)
Male	9 (40.9%)	146 (36.3%)	246 (40.7%)	67 (41.6%)
Race				
Caucasian	16 (72.7%)	329 (81.8%)	468 (77.5%)	121 (75.2%)
African American	2 (9.1%)	21 (5.2%)	46 (7.6%)	6 (3.7%)
Asian	2 (9.1%)	23 (5.7%)	33 (5.5%)	12 (7.5%)
Hispanic/Latino	1 (4.5%)	18 (4.5%)	24 (4.0%)	16 (9.9%)
Other	1 (4.5%)	11 (2.7%)	33 (5.4%)	6 (3.7%)
Sexual Orientation				
Heterosexual	19 (86.4%)	374 (93.0%)	556 (92.1%)	143 (88.8%)
Bisexual	3 (13.6%)	14 (3.5%)	28 (4.6%)	10 (6.2%)
Lesbian/Gay	--	12 (3.0%)	14 (2.3%)	5 (3.1%)
Other	--	2 (0.5%)	6 (1.0%)	3 (1.9%)
Marital Status				
Single	17 (77.3%)	267 (66.4%)	449 (74.3%)	133 (82.6%)
Married	4 (18.2%)	113 (28.1%)	117 (19.4%)	25 (15.5%)
Separated	1 (4.5%)	8 (2.0%)	6 (1.0%)	--
Divorced/Widowed	--	14 (3.4%)	32 (5.3%)	3 (1.9%)
Employment Status				
Unemployed	7 (31.8%)	192 (47.8%)	257 (42.5%)	71 (44.1%)
Part time	7 (31.8%)	100 (24.9%)	220 (36.4%)	49 (30.4%)
Full time	8 (36.4%)	110 (27.4%)	127 (21.0%)	41 (25.5%)

^aAge range for community group was 24–69 in Study 1 ($M = 40.30$, $SD = 13.61$), 20–70 in Study 3 ($M = 38.56$, $SD = 12.76$), 18–74 in Study 4 ($M = 36.42$, $SD = 12.42$), and 20–69 in Study 5 ($M = 33.59$, $SD = 8.97$)

^bAge range for student group was 18–21 in Study 1 ($M = 19.33$, $SD = 0.78$), 18–62 in Study 3 ($M = 19.83$, $SD = 4.61$), 18–50 in Study 4 ($M = 19.48$, $SD = 3.06$), and 18–29 in Study 5 ($M = 19.34$, $SD = 1.59$)

Table 2

PIES-P Item Means and Factor Weights (Study 3) and PIES Standardized Factor Loadings (Study 4)

PIES-P Items (Study 3)	Study 3 Mean (SD)	Study 3 Factor Weights (PAF)	Retained in PIES	Study 4 Standardized Factor Loadings ⁷
When I share how I'm feeling, others...				
1. ...act like I have no reason to be upset.	1.91 (0.94)	.74		
2. ...don't seem to mirror or match my emotions. For example, they don't share sadness with me when I'm sad or happiness with me when I'm happy.	1.98 (1.00)	.67	✓	.68
3. ...act like they don't care.	1.76 (0.99)	.79		
4. ...don't seem to understand why I feel the way that I do.	2.03 (1.06)	.78		
5. ...act like things are not that bad.	2.22 (1.05)	.76		
6. ...try to change how I feel rather than just understand me.	2.24 (1.07)	.65		
7. ...try to fix my problems without understanding how I'm feeling.	2.20 (1.04)	.60		
8. ...blame me for feeling the way that I do.	1.73 (1.01)	.79		
9. ...want me to "get over it" or "accept it and move on."	2.19 (1.12)	.73	✓	.71
10. ...seem like they don't want to hear what I have to say.	1.91 (1.07)	.81	✓	.85
11. ...look down on me or judge me.	1.68 (0.96)	.79	✓	.86
12. ...act like I'm blowing things out of proportion.	2.07 (1.03)	.75		
13. ...make it all about themselves rather than just take the time to listen to me.	2.05 (1.12)	.74		
14. ...don't take me seriously.	1.78 (0.99)	.82	✓	.83
15. ...express disapproval or disappointment.	1.65 (0.97)	.80		
16. ...tell me or imply what I should actually feel.	1.94 (1.05)	.76	✓	.72
People...				
17. ...around me make me feel like a robot because I don't show enough emotion, or like a crybaby because I'm too emotional	1.87 (1.16)	.59		
18. ... say that I'm only feeling the way that I do because of who I am. For example, by saying, "Well you just feel that way because you're ____ (a man/a woman/liberal/young/etc.)__!"	1.73 (0.98)	.69		
19. ... say that I'm not feeling what I should because of who I am. For example, by saying, "The only reason don't feel that way is because you're ____ (a man/a woman/liberal/young/etc.)__!"	1.66 (0.97)	.65		
Others...				
20. ... tell me that they understand how I feel even though I know that they don't.	2.23 (1.07)	.66		
21. ... pick my feelings apart from every angle.	1.60 (0.92)	.67		
22. ...get mad or upset at me when I express my feelings.	1.63 (0.97)	.79	✓	.80
23. ...don't take my side or agree with how I'm feeling.	1.97 (1.03)	.81	✓	.79
24. ...make me feel like it's not okay for me to feel the way that I do.	1.84 (1.07)	.76	✓	.84
25. ...make me feel guilty about my emotions.	1.76 (0.99)	.77		
26. ...make me feel that my emotions are unimportant.	1.80 (1.05)	.82	✓	.79

⁷Factor loadings were computed based on dataset 2.

PIES-P Items (Study 3)	Study 3 Mean (SD)	Study 3 Factor Weights (PAF)	Retained in PIES	Study 4 Standardized Factor Loadings⁷
27. ...act like my emotions don't make sense.	1.87 (1.07)	.82		

⁷Factor loadings were computed based on dataset 2.

Bivariate Correlations for PIES and Study 5 Variables Measured at Time 1 and Time 2

Table 3

Construct (Scale)	1	2	3	4	5	6	7
1. Current Emotion Invalidation (PIES)	--	.59**	.46**	-.36**	-.47**	-.45**	-.42**
2. Emotional Distress (DASS-21)	.45**	--	.65**	-.50**	-.63**	-.52**	-.41**
3. Emotion Dysregulation (DERS)	.47**	.56**	--	-.47**	-.58**	-.45**	-.29**
4. Physical Health (WHOQOL)	-.28**	-.58**	-.46**	--	.69**	.48**	.60**
5. Psychological Health (WHOQOL)	-.46**	-.62**	-.59**	.67**	--	.59**	.64**
6. Relational Health (WHOQOL)	-.35**	-.49**	-.39**	.52**	.64**	--	.46**
7. Environmental Health (WHOQOL)	-.41**	-.43**	-.35**	.67**	.64**	.51**	--
Mean at Time 1	1.69	22.10	2.14	16.87	14.91	14.93	15.57
SD at Time 1	0.66	18.75	0.61	2.43	2.85	3.51	2.49
Mean at Time 2	1.80	18.61	2.10	16.87	14.96	14.89	15.56
SD at Time 2	0.72	19.24	0.60	2.63	3.00	3.39	2.67

Note: Correlations between Time 1 variables appear above the diagonal, and correlations between Time 2 variables below.

**

 $p < .01$

Table 4

Three Hierarchical Regressions Predicting Emotion Dysregulation (DERS Total Scores), Borderline Features (MSI-BPD) and Emotional Distress (DASS-21)

	DERS Total Scores	MSI-BPD	DASS-21
	B (SE)	B (SE)	B (SE)
Step 1	($R^2 = .19^{**}$)	($R^2 = .10^{**}$)	($R^2 = .12^{**}$)
Sample ^a	-.43 (.09)**	-.74 (.40)	-5.65 (2.84) ⁺
Social Desirability (MCSF)	-.06 (.02)**	-.23 (.08)**	-1.76 (.59)**
Childhood Invalidation (SES)	.13 (.05)**	.51 (.20)*	4.34 (1.43)**
Step 2	($R^2 = .12^{**}$)	($R^2 = .08^{**}$)	($R^2 = .25^{**}$)
Sample	-.36 (.08)**	-.50 (.39)	-2.51 (2.44)
Social Desirability (MCSF)	-.04 (.02)*	-.16 (.08) ⁺	-.86 (.52)
Childhood Invalidation (SES)	.07 (.04)	.29 (.20)	1.55 (1.26)
Current Invalidation (PIES) ^b	.34 (.07)**	1.18 (.31)**	15.23 (1.93)**

Notes:

^a Student participants were coded as 0 and MTurk workers were coded as 1 for these analyses.

^b Scores at Time 1 administration.

⁺ $p = .05$,

* $p < .05$,

** $p < .01$

Table 5

Three Hierarchical Regressions Predicting Emotional Distress (DASS-21) and Relational and Psychological Health (WHOQOL subscales)

	Emotional Distress	Relational Health	Psychological Health
	B (SE)	B (SE)	B (SE)
Step 1	($R^2 = .55^{**}$)	($R^2 = .57^{**}$)	($R^2 = .65^{**}$)
Time 1 Scores on Corresponding Outcome Measure ^a	.76 (.06) ^{**}	.73 (.05) ^{**}	.85 (.05) ^{**}
Step 2	($R^2 = .01^+$)	($R^2 = .03^{**}$)	($R^2 = .02^{**}$)
Current Invalidation (PIES) ^b	3.19 (1.91) ⁺	-.92 (.29) ^{**}	-.72 (.23) ^{**}

Notes:

^aDASS-21 scores at Time 1 were entered in Step 1 for emotional distress outcome, WHOQOL scores for relational and psychological health at Time 1 were entered in Step 1 for relational and psychological health outcomes respectively.

^bScores at Time 2 administration.

⁺ $p = .10$,

^{*} $p < .05$,

^{**} $p < .01$