Behavioral assessment of the negative emotion aspect of distress tolerance: Tolerance to emotional images

Jennifer C. Veilleux[1]

Garrett A. Pollert[1]

Melissa J. Zielinski[2]

Jennifer A. Shaver[3]

Morgan A. Hill[1]

[1]University of Arkansas, Fayetteville

[2]University of Arkansas for Medical Sciences

[3]University of Washington, Seattle

Correspondence should be addressed to:

Jennifer C. Veilleux
216 Memorial Hall
Department of Psychological Science
University of Arkansas
Fayetteville, AR, 72701
jcveille@uark.edu
Phone: 479-575-5329
Fax: 479-575-3219

Abstract

The current behavioral tasks assessing distress tolerance measure tolerance to frustration and tolerance to physical discomfort, but do not explicitly assess tolerance to negative emotion. We closely evaluated the conceptual distinctions between current behavioral tasks and self-report tasks assessing distress tolerance, and then developed a new behavioral distress tolerance task called the Emotional Image Tolerance (EIT) task. The EIT task retains elements of existing behavioral tasks (e.g., indices of persistence) while augmenting the reliability and content sufficiency of existing measures by including multiple trials, including a variety of negative affect stimuli, and separating overall task persistence from task persistence after onset of distress. In a series of three studies, we found that the EIT correlated with extant behavioral measures of distress tolerance, the computerized mirror tracing task and a physical cold pressor task. Across all of the studies, we also evaluated whether the EIT correlated with self-report measures of distress tolerance and measures of psychopathology (e.g., depression, anxiety, and binge eating). Implications for the refinement of the distress tolerance construct are discussed.

Keywords: Distress tolerance; distress intolerance; behavior; self-report; measurement

There is significant evidence that self-report measures and behavioral tasks assessing distress tolerance are only modestly correlated, and often not at a statistically significant level (Ameral, Reed, Cameron, & Armstrong, 2014; Berstein, Marshall, & Zvolensky, 2011; Cougle, Bernstein, Zvolensky, Vujanovic, & Macatee, 2013; Kiselica, Rojas, Bornovalova, & Dube, 2015). Self-report measures capture the *perceived capacity* of withstanding aversive emotional and physical states whereas the behavioral tasks capture the *actual behaviors* associated with withstanding these aversive states (Cougle et al., 2013; Glassman et al., 2016; Leyro, Zvolensky, & Bernstein, 2010). Of course, deviations between self-perceptions of behavior and actual behavior are neither new nor unique to distress tolerance; discrepancies between self-report, informant report, implicit measures and behavioral measures are rampant in many areas of psychology (Mischel, 1968; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). In most cases, it seems reasonable to think that self-perceptions and behavioral abilities are capturing different aspects of the same latent construct.

Distress tolerance is a bit different. Distress tolerance is an umbrella term used to describe several lower-order specific constructs: tolerance of uncertainty, tolerance of ambiguity, tolerance of frustration, tolerance of negative emotion, and tolerance of physical discomfort (Zvolensky, Vujanovic, Bernstein, & Leyro, 2010). All five of these lower-order aspects of distress tolerance have been assessed via self-report (see Leyro et al., 2010 for review). However, only two of these lower-order constructs have been clearly assessed by behavioral tasks. Specifically, the Paced Serial Addition Test (PASAT; e.g., Lejuez, Kahler, & Brown, 2003) and the mirror tracing task (e.g., Strong et al., 2003) assess tolerance of frustration, and the other commonly used tasks such as the cold pressor task (e.g., Daughters, Lejuez, Bornovalova, Kahler, Strong, & Brown, 2005) or breath holding tasks (e.g., Daughters et al., 2005; Zvolensky, Feldner, Eifert, & Brown, 2001) assess tolerance of physical discomfort. It is plausible that physical pain or discomfort are associated with negative emotion, and frustration is in certainly in the "anger family" (Ekman & Cordaro, 2011). But, none of the extant tasks capture tolerance to sadness, disgust, or anxiety—emotions that are considered the other "basic"

negative emotions (Ekman & Cordaro, 2011). Recent work has moved in this direction, such as by assessing momentary self-reported tolerance to negatively valanced film clips (Cougle et al., 2013), though that still relied on self-report.

These two separate issues, the limited concordance between self-report and behavioral measures and the need to develop behavioral measures to assess other lower-order aspects of distress tolerance, may be linked. For example, one could predict that the modest relationships between self-report and behavioral indices occur because studies use measures that assess different lower-order aspects of distress tolerance. There is little evidence of this, however, as measures within the same domain (e.g., a behavioral task of physical tolerance and a self-report measure of physical tolerance) also show modest (Kiselica, Rojas, Bornovalova, & Dube, 2015), often non-significant correlations (Bernstein, Marshall, & Zvolensky, 2011; Glassman et al., 2016).

Another approach could be to hone in on one lower-order component, such as tolerance to negative emotion, and intentionally develop a behavioral measure to assess tolerance to negative emotion that maps closely onto the self-report measures assessing tolerance to negative emotion. This would serve multiple functions, the first of which is expansion of the measurement of distress tolerance, such that behaviorally-indexed tolerance to negative emotion may have different correlates and predictors than frustration or pain tolerance. For example, we know that substance users (Daughters et al., 2005) and people with borderline personality disorder (Bornovalova, Gratz, Daughters, Nick, Delany-Brumsey, & Lynch, 2008) show lower tolerance to frustrating and physical discomfort tasks. Perhaps tolerance to negative emotion may be more salient for internalizing disorders such as anxiety and depression. Prior research has shown significant correlations between symptoms of internalizing psychopathology and self-report measures of tolerance to negative emotion (Anestis et al, 2012; Bernstein et al, 2011). Without a measure to behaviorally assess tolerance to negative emotion we are stunted in our understanding of the role of distress tolerance as a hypothesized transdiagnostic risk factor (Zvolensky et al., 2010).

The second function of trying to develop a new behavioral task to map onto a self-report measure involves explicitly examining the "method-construct confound" (Leyro et al., 2010, pp. 591). The purpose of this isn't to try to "force" behavioral tasks and self-report measures to correlate, rather to try to more deeply understand the ways in which measures that purportedly assess the same construct—in particular, the same lower-order construct—differ at a conceptual level, with the goal of informing the overall distress tolerance construct (Ameral et al., 2014; Glassman et al., 2016; Kiselica et al., 2015; McHugh & Otto, 2011).

Conceptual Distinctions Based on Measurement Method

There are multiple reasons why self-report measures of frustration tolerance and behavioral frustration tolerance tasks are distinct, which reflect both methodological and conceptual distinctions. To discuss these issues, it may be fruitful to look at distinctions between existing measures in the same lower-order domains such as between the Frustration Discomfort Scale (FDS; Harrington, 2005) and the two behavioral frustration tolerance tasks, the PASAT (Lejuez et al., 2003) and the mirror tracing task (e.g., Strong et al., 2003).

First, unlike the self-report measures, the behavioral tasks are typically goal-oriented persistence tasks. People must try to complete the backward trace of a figure in the mirror tracing task (Strong et al., 2003) or correctly add successive numbers in the PASAT (Lejuez et al., 2003), and the task assesses how long a person persists until they "give up." In addition, these tasks are also often (Bornovolova et al., 2008; Brown, Lejuez, Kahler, & Strong, 2002) though not always (Feldman, Dunn, Stemke, Bell, ^ Greeson, 2014; Macatee & Cougle, 2015) administered with financial incentives to try to increase effort on the task. The self-report measures, on the other hand, rarely address persistence and don't involve financial incentives. A few items on the FDS address task-related behavior, such as "I can't stand doing tasks when I'm not in the mood," but although these items suggest an unwillingness to persist on tasks when upset, they do not assess persistence directly.

Instead, many items on the FDS ask about tolerating the *experience* of distress (e.g., "I can't bear disturbing feelings"), implying that the conscious feeling of distress is what cannot be tolerated. Yet, the current persistence tasks do not necessarily create conscious distress. At the group level, pre- to post-task ratings indicate that both the PASAT and the mirror tracing task are associated with increased frustration (Ameral et al., 2014). However, these kind of analyses don't ensure that frustration is experienced at the individual level. In fact, a recent study asked participants to narratively describe why they stopped persisting on the PASAT and mirror tracing tasks (Ameral et al., 2014) and only about half of participants described feeling upset or bothered by the tasks as a reason for task termination. From the perspective of the self-report measures, the idea is that before distress can be tolerated, it must be present; persisting on a task that does not induce distress would not truly index distress tolerance. A new behavioral task that adopted the perspective of the self-report scales would thus attempt to firmly establish distress for each person prior to assessing how long a person is willing to endure or persist on that task.

Third, the FDS—and all self-report measures of distress tolerance—addresses perceptions of the self, which requires people to address *how they are generally*, rather than how they might behave in a specific moment. To provide a rating to the FDS questions, people have to think about how they typically behave over time and across many situations. The behavioral tasks, on the other hand, assess ability in one situation, at one moment in time; they assess "snapshots" of behavior. It is theoretically reasonable to use these behavioral "snapshots" as indicative of a person's typical behavior, since a person with "low distress tolerance" should show low persistence in multiple contexts. However, a long history of research in social psychology suggests that behavior is variable across time and situation situations (e.g, Fleeson, 2001; Mischel, 1968; Ross & Nisbett, 1991). Thus, even if we adopt the assumption that we can accurately report on our own behavior (which is almost certainly a flawed assumption; Bem, 1972; Nisbett & Wilson, 1977), any report or assessment of *general* or *average* behavior is likely to be different than a report or assessment of a *single* behavior.

Finally, and related to the previous point, there are psychometric distinctions between the self-report and behavioral tasks. For example, the FDS has multiple items; scores on 35 individual items are averaged for a total score, or grouped into 7-item subscales. Development of self-report scales involves psychometric analysis of internal consistency and reliability of the items (Kiselica, Webber, & Bornovalova, 2014) to assess how well the items "hang together" as indicators of one latent construct. An FDS total score should be, statistically speaking, relatively robust against measurement error or variability of individual items. Yet, most of the behavioral tasks—including the PASAT and mirror tracing task—are single item measures, which makes them more susceptible to contextual influences, including measurement error and situational factors. A behavioral task that has more trials and provides an average score would be statistically more reliable. It would not entirely address the conceptual issue, because multiple trials completed in a lab still assess the same behavior in relatively the same context, but if the trials varied in *type*, an average score could more closely approximate average or *general* behavior.

New Task Development

The current set of studies was designed to develop a new behavioral distress tolerance task assessing tolerance to negative emotion. We wanted a task that used a persistence-based measure to allow for some parity between the new task and extant behavioral measures, but we also wanted to explicitly narrow the conceptual gaps between our task and commonly used self-report measures of tolerance to negative emotion such as the Distress Tolerance Scale (DTS; Simons & Gaher, 2005).

We therefore devised a task which required individuals to tolerate negative emotional images. We chose images due to the decades of research using images to evoke negative mood (Coan & Allen, 2007; Ellard et al., 2012), where existing and validated images cover a variety of negative emotions (e.g., disgust, fear, sadness, anger). We also thought it feasible for one participant to view multiple images, thus allowing for elicitation of broad negative mood where multiple trials would allow analysis of variability and internal consistency. In addition, we developed a novel approach for ensuring that

participants experienced distress before having to tolerate distress. We wanted to be able to separate overall persistence, operationalized as the amount of time a person views a negative image, from *distress tolerance*, which theoretically should be defined as the amount of time a person is willing to sit with a negative image *after feeling distressed by it.* The overall persistence concept is similar to existing behavioral assessments, and the latter comes closer to how distress tolerance is conceptualized in the DTS.

We report three studies conducted to develop this new task. First, we empirically selected images for the task and conducted preliminary comparisons between our task and a self-report measure of tolerance to negative emotion (Study 1). We predicted that our task would show significant but low to moderate correlations with the self-report measure; significant because of our efforts to map the new task onto the self-report measure, but low because of inherent differences between self-perceptions and observable behaviors. Second, we wanted to examine our task against existing behavioral measures from two other lower-order domains of distress tolerance. To do so, we evaluated our task alongside the mirror tracing test (Study 2) and e the cold pressor task (Study 3). We anticipated that our new task would correlate significantly with extant persistence tasks but also at a low magnitude, as the tasks assess different lower-order domains of distress tolerance. Finally, in Studies 2 and 3 we also evaluated our new task alongside additional self-report measures and with several indices of psychopathology in an initial attempt to evaluate how behaviorally indexed tolerance to negative emotion might relate to both internalizing and externalizing pathology.

Study 1

Method

Participants. Participants were recruited through a mid-southern university subject pool, and received course credit for participation. Recruitment occurred via an online screener. Due to the graphic nature of the images and associated concerns by our university Institutional Review Board, we attempted to minimize risk by excluding participants who were experiencing heightened psychological

symptoms. This was particularly important for Study 1, where participants saw more images than in the subsequent studies. We excluded participants with Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983) scores more than 1.5 standard deviations above the mean for college students of the same reported gender (above 1.66 for men and above 1.34 for women; Cochran & Hale, 1985). We also excluded participants who self-reported currently taking psychotropic medications or who had participated in psychotherapy within the last six months. Finally, we excluded participants who self-reported diagnoses of social anxiety, PTSD or panic disorder, as we anticipated that people with these diagnoses might have particularly heightened physiological reactions to the images. In total, 119 participants were screened, and 27 were deemed ineligible based on the criteria above. Significantly more of the ineligible group were women (96%), $\chi 2 = 8.81$, $p = .003$. Notably, our exclusion strategy was used to minimize risk by excluding the most vulnerable participants who self-reported current issues. We recognize this strategy did not truly ensure a "healthy" sample; we hoped to retain variability in responses to the images in a broad sample of college students. Eligible participants were invited to sign up for laboratory session slots, and 48 participated in the current study (68.8% female, 79.2% Caucasian, mean age = 20.00).

Measures.

Affect. State positive affect (PA) and negative affect (NA) were measured with the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), a 20-item measure with ten adjectives for each scale, where participants indicate how they are feeling right now, at this moment, from 1 (*very slightly or not at all*) to 5 (*extremely*). The positive affect subscale demonstrated adequate reliability, with alphas at .89 to .91 across the two time points. Negative affect had an alpha of .60 at the initial time point and .80 at the post-task time point.

Self-Reported Distress Tolerance. Self-reported perceptions of distress tolerance were measured using the Distress Tolerance Scale (DTS; Simons & Gaher, 2005), a 14-item index of the ability to tolerate general distress. Items are measured on a 1 (*strongly agree*) to 5 (*strongly disagree*) Likert-type

scale, where higher scores represent greater perceptions of distress endurance. The DTS had excellent internal consistency ($\alpha = .93$).

Behavioral Distress Tolerance. Participants were told they would be viewing pictures on a computer screen and responding via keyboard entry. They were instructed to press "q" at the first sign of distress or discomfort with each image and to continue viewing the image until the discomfort was too much to handle, at which point they could press "p" to move on to the next picture. For this initial iteration, distress (i.e., pressing the letter "q") had to be noted within the first 15 seconds. If distress was not indicated within the first 15 seconds, the program automatically moved to the next image. If distress *was* noted within the first 15 seconds, the participant was given 15 additional seconds to tolerate the image after indicating distress. If after denoting distress, an image was deemed too difficult to tolerate, a person could press "p" to advance to the next image, any time during the "tolerating" portion of the slide. Therefore, for Study 1 each slide was viewed from between 15 seconds (for images that were not distressing) to just under 30 seconds (if a person indicated distress at about the 15-second mark, and did not select "p" during the latter 15 seconds). The images were 59 negative slides from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) selected after initial pilot testing of 80 slides[1].The slides were presented using EPrime 2.0 on a 22" monitor. Slides were presented in five blocks of 11 or 12 slides each, with slides randomized to order within block and blocks presented in random order.

[1] Participants in the pilot (n = 30, 46.7% female, 90% Caucasian, mean age 19.33 (SD = 1.45) viewed 80 slides in random order for up to 15 seconds each and asked to press "p" when they felt distress or discomfort (if they did not denote distress, the next slide came up after 15 seconds). After the initial presentation, participants viewed each slide a second time and rated their degree of distress, anger, anxiety, disgust, sadness, fear and happiness on 0 (not at all) to 6 (strongly) Likert-type scales. We removed 12 slides with mean "p" time (i.e., time to denote distress) of above 10 seconds, and we excluded an additional 9 slides with emotion distress ratings of less than 2.67 (i.e., more than one standard deviation below the mean distress rating). This resulted in exclusion of 21 slides that were less distressing than the retained slides, both in terms of time to denote distress (M = 5.64 for retained slides, M = 10.07 for excluded slides, $t(78) = 7.96$, $p < .001$) and ratings of distress (M = 3.74 for retained slides, M = 2.40 for excluded slides, $t(78) = 7.35$, $p < .001$).

Slide Distress Ratings. After completion of the persistence task, all images were viewed again and rated on a 1 (*no distress)* to 6 (*extreme distress*) Likert-type scale for the degree of distress elicited by the slide.

Procedure. Following consent, participants completed the DTS and an initial PANAS to assess baseline mood. The experimenter then introduced the participant to the behavioral distress tolerance task, walked the participant through two examples, verbally confirmed understanding of the task, and then had the participants view all 59 slides. The experimenter stayed in the room during the task, both to ensure the participant continued to look at the screen, and in case any participant found the images too distressing and wanted to drop out of the study. After the task, participants then completed a second PANAS, and rated all slides for degree of distress. After the distress ratings, participants were asked to view a humorous video of about 5 minutes in length (they were allowed to choose from a blooper reel of a sitcom or a funny cat video), and were thoroughly debriefed before leaving the lab. Experimenters probed for prolonged distress and any participant who appeared to have lingering reactions to the slides was asked to view the other humorous video.

Results

We first examined the set of images used in the persistence task with the intention of reducing the image set to 45 slides. We aimed for 45 slides to allow for adequate emotional content sampling and assessing of reliability, yet short enough to be completed within 25 minutes, even for participants who viewed each slide for the maximum amount of time (30 seconds). Because we wanted to retain slides that were considered distressing by most participants, and discard slides that were too easy to tolerate, we calculated the percent of participants who found each slide distressing (e.g., people who hit "q" at some point during the first 15 seconds), and the percentage of participants who viewed each slide for the entire time (see Table 1 for descriptive statistics on all 59 slides). Additional image statistics are also listed in Table 1, namely the average distress threshold (e.g., average time into the slide participants hit "q"), distress persistence (e.g., after selecting "q," the average time participants viewed the slide before

moving on to the next slide), and the total time each slide was viewed, as well as the average distress rating given to each slide in the post-task emotion ratings. We used the latter to discard the 14 slides with the lowest distress ratings, which included the 5 slides with distress threshold ratings above 10 seconds and the majority of slides where more than 50% of participants viewed the slide for the entire possible viewing time. We call this final set of 45 images and the associated directions the Emotional Image Tolerance (EIT) task.

After selecting the top 45 slides, we calculated two variables from the EIT data for each participant. Distress threshold was calculated by averaging the initial distress time (i.e., time to pressing "q") across the 45 slides for each participant; reliability on this index was extremely high ($\alpha = .98$). Distress persistence was calculated by computing the average amount of time the participant spent looking at the slide *after* distress was indicated; by necessity, this value only included slides where distress was reported. Persistence reliability across the 45 slides was also very high ($\alpha = .99$).

Paired samples *t*-tests revealed that participants increased in negative affect (NA) from pre- ($M = 12.23$, $SD = 2.42$) to post-task ($M = 16.19$, $SD = 2.48$), $t(46) = 5.86$, $p < .001$. Similar declines in positive affect (PA) were found from pre- ($M = 26.43$, $SD = 7.72$) to post-task ($M = 21.13$, $SD = 7.60$), $t(46) = 7.15$, $p < .001$.

As an initial test of task validity, we evaluated zero-order correlations between the EIT measures, self-reported distress tolerance, average distress ratings across the slides from the post-task re-evaluation of each image, and affect (see Table 2). To capture affect change, we used post-task affect scores controlling for pre-task affect scores via partial correlations. We found evidence that higher post-task NA, controlling for pre-task NA, was associated with lower EIT distress threshold. PA was not associated with any of the distress tolerance variables. We also found that higher self-reported distress tolerance was significantly correlated with longer distress persistence. Finally, people who rated the slides as more distressing had lower distress thresholds and lower distress persistence. They also had higher post-task NA and lower post-task PA.

Discussion

The intent of Study 1 was to develop and provide initial validation for a behavioral measure of tolerance to negative emotion. The EIT evidenced several strengths. Namely, unlike the majority of prior work (Ameral et al., 2014; Anestis et al., 2012; Bernstein et al., 2011; Glassman et al., 2016; McHugh, Daughter et al., 2011), the EIT demonstrated a small but significant relationship with self-reported distress tolerance, as predicted. We would like to assume that these significant relationships are due to our choices when developing the task as intentionally closer to the self-report measures. Specifically, because we chose to include multiple trials, we were able to establish reliability of the image responses. Because we included a variety of emotionally evocative negative images, we theoretically captured a broader negative emotional profile. Finally, our calculation of a persistence index ensures—at least at a face valid level—that participants persisted through actual experienced distress. Not only did the distress persistence measure capture felt distress, but people felt worse (greater NA) after the EIT than before the EIT. Post-task NA was associated with distress threshold, whereas self-reported distress tolerance showed no relationship with affect. Of note, these results were found with a task that uses a similar assessment of persistence as extant behavioral tasks—time until "giving up" or ceasing effort, similar to the mirror tracing task and the cold pressor (Leyro et al., 2010), albeit without financial incentives or a clear completion goal.

Study 2

The primary function of this study was to compare the newly developed EIT alongside the mirror tracing task. The secondary function was to examine the relationships of both behavioral tasks to self-report measures of distress tolerance and to symptoms of psychopathology (Bernstein et al., 2011; Kiselica et al., 2015). Although past work with behavioral tasks has targeted substance users (Daughters Sargeant, Bornovolova, Gratz, & Lejuez, 2008; Kiselica et al., 2015) and borderline personality disorder (Bornovalova et al., 2008) or borderline features (Kiselica et al., 2014), due to our college sample we chose to focus on more common indices of pathology. Specifically, we assessed symptoms of

depression, anxiety and problematic eating behavior, which are all associated with lower self-reported distress tolerance (Anestis et al., 2012; Bernstein et al., 2011).

Method

Participants & Recruitment. The same recruitment procedures were used as in Study 1. In total, 739 participants were screened, and 612 were eligible, with no gender differences in percentages of eligible participants. Eligible participants were invited to sign up for laboratory session slots, and 73 participated in the current study (65.8% female, 80.80% Caucasian, mean age = 19.05).

Measures.

Affect. The PANAS (Watson et al., 1988) was used to assess affect throughout the study. The PANAS was given at baseline and after each behavioral task, and we only focus on negative affect in our analyses below. Internal consistency for negative affect ranged between .70 and .87.

Self-Reported Distress Tolerance. Self-reported perceptions of emotional distress tolerance were measured using the Distress Tolerance Scale (DTS; Simons & Gaher, 2005), where the alpha for the current study was .90.

We also used the Discomfort Intolerance Scale (DIS; Schmidt, Richey, & Fitzpatrick, 2006), a 5-item self-report measure of difficulties withstanding uncomfortable physical sensations, including pain (e.g., I take extreme measures to avoid feeling physically uncomfortable). The DIS items are rated on a 6-point Likert-type scale from 0 (*not at all like me*) to 6 (*extremely like me*). The scale has demonstrated high internal consistency in addition to both convergent and divergent validity with regards to similar constructs (e.g., anxiety sensitivity; Schmidt et al., 2006). In the current study, the alpha was .68.

Psychopathology. We used the Binge Eating Scale (BES; Gormally, Black, Datson & Rardin, 1982), a 16-item measure, to assess binge-eating severity. Responses reflect the participant's behavioral, cognitive, and emotional experiences surrounding a binge episode. The items are rated on a 3-point

Likert-type scale from 1 (*no binge eating problems*) to 3 (*severe binge eating problems*), and alpha in the current study was acceptable at .85.

We also assessed current mood symptoms with the Depression Anxiety and Stress Scales (DASS-21; Henry & Crawford, 2005), a 21-item measure measuring anxiety, depression, and stress and is a shortened version of the original 42-item self-report measure of depression, anxiety, and stress (DASS; Lovibond & Lovibond, 1995). The items are measured on a 0 *(did not apply to me at all)* to 3 *(applied to me very much, or most of the time*) scale. We used the total score which reflects the degree to which an individual is experiencing general psychological distress, where the alpha in the current study was .90.

Behavioral Distress Tolerance.

Mirror Tracing Persistence Task-Computerized (MTPT-C; Strong et al., 2003). In the MTPT-C, participants use a computer mouse to "trace" the outline of a figure backward. Whenever a participant deviates from the figure, the computer produces a loud error buzzing noise, and the participant must start the task again. After completing two relatively simple practice figures, the participant is introduced to a difficult figure, and told they can stop the task at any time by pressing a keyboard key. This task has been used in many studies as a behavioral index of distress tolerance (Ameral et al., 2014; Bernstein et al., 2011; Bornovolova et al., 2008; Feldman et al., 2014; Macatee & Cougle, 2015), and just as in several prior cases (Feldman et al., 2014; Macatee & Cougle, 2015), no incentives were provided for longer persistence. In addition to the task itself, the MTPT-C includes brief measures of irritability, frustration, anxiety, difficulty concentrating, and bodily discomfort on 0 (*none*) to 100 (*extreme*) visual analogue scales.

Emotional Image Task (EIT). The EIT was administered here with slightly different programming than in Study 1. Because this is the version of the task used in subsequent studies, instructions and processes are described here in full. Each slide could be viewed for a maximum of 30 seconds. Participants were told they would be viewing pictures on a computer screen and responding

via keyboard entry. They were instructed to press "q" at the first sign of distress or discomfort with each image and to continue viewing the image until the discomfort was too much to handle, at which point they could hit "p" to move on to the next picture. Of note, the participant could not move on from the slide (by hitting "p") without first denoting distress (by hitting "q"). Any slide that was not denoted as distressing was viewed for the entire 30 seconds. The images were 45 negative images selected from the IAPS (Lang, Bradley, & Cuthbert, 2008), presented using EPrime 2.0 on a 22" monitor. Slides were presented in 5 blocks of 9 slides each, where slides similar in content were placed into different blocks. Slides were randomized to order within block and blocks presented in random order; these procedures attempted to ensure that similar pictures would not appear next to one another.

Four variables were calculated for each person. First, the average amount of time the participant viewed each slide was calculated as the mean viewing time across the set of 45 slides (with a possible range of 0 to 30 seconds), which we call *image persistence*. Second, the average *distress threshold* was calculated as the average time into the slides the participant acknowledged distress, or the average time until the participant pressed the letter "q" on the keyboard (possible range of 0 to 30 seconds). For this rating, any slides that were not noted as distressing (e.g., "q" was never pressed) were given a value of 30 seconds. Third, we calculated a count of how many slides (from 0 to 45) on which the participant noted any distress. Finally, we calculated *distress persistence*, operationalized as the time *after* a participant noted distress before they moved on to the next slide, or the average time between the participant pressing "q" and pressing "p." This measure by necessity only includes slides that a person actually found distressing.

Procedure. After informed consent, participants were told they would be completing two "mini-studies." They began by completing individual difference measures, followed by the first administration of the PANAS to assess baseline mood. They were then randomly assigned to order of tasks, and completed either the computerized version of the mirror tracing task, or the EIT. After first task, participants completed a second PANAS, and then the opposite task as part of the other "mini-

study." Finally, participants completed a third PANAS. Just as in Study 1, participants watched a funny video to repair mood prior to debrief.

Results

Data preparation. We evaluated total time on the mirror tracing task as one outcome variable, where higher scores represent longer persistence on the task. We also created an index of mirror tracing skill by taking the number of times the participant returned to the start (e.g., number of errors) divided by the total task time, to arrive at an index of errors per second (Bornovolova et al., 2008; Feldman et al., 2014; MacPherson, Stipelman, Duplinsky, Brown, & Lejuez, 2008). With this index, higher scores represent more errors per second and thus less skill at the mirror tracing task. Due to significant skewness and kurtosis of both measures, the variable were log transformed prior to analysis, and all subsequent analyses involving mirror tracing use the log transformed variables. None of the four EIT variables were significantly skewed or kurtotic.

Because participants completed the tasks in randomized order, we tested for order effects. Using an independent samples *t*-test, we found that people who completed the mirror tracing first had more errors per second (lower skill; M = -.05, SD = .32) than people who completed the EIT first (M = -.26, SD = .33), $t(70) = -2.78$, $p < .01$, although there were no differences between groups on mirror tracing completion time, $t(70) = .51$, $p = .61$. We also found no order effects on any of the four EIT measures. To account for the order effect, all analyses involving the mirror tracing skill controlled for order.

Examining EIT Descriptives. Because the EIT is a new task, we wanted to examine the descriptive statistics of the four EIT variables, to get a sense of how participants responded to the task. The average image viewing time was 16.30 seconds (SD = 9.74, range .68 to 30). There were 7 participants who viewed all 45 slides for the entire 30 seconds. Participants indicated distress (e.g., pressed "q") on 33.68 (SD = 13.87) of the 45 slides, where only 1 person found no slides distressing, and 22 people found all 45 slides distressing. Distress threshold was, on average, 10.53 seconds into the slide (SD = 9.05), and participants persisted in viewing the slide for an average of 8.75 seconds (SD =

7.15) after denoting distress (range .11 to 24.47 seconds). All of the EIT variables are correlated with each other (see Table 3).

Affect. We compared negative affect after each task and baseline negative affect using within-subject ANOVA, controlling for order, $F(2, 138) = 17.56$, $p < .001$. We found that post-EIT negative affect was significantly higher ($M = 17.34$, $SD = 6.75$) than post-mirror tracing negative affect ($M = 13.96$, $SD = 3.85$), which in turn was higher than baseline negative affect ($M = 11.93$, $SD = 2.58$), all Bonferroni-corrected post-hoc tests were significant at $p < .001$. All subsequent analyses involving affect use post-task negative affect controlling for baseline negative affect to provide an indication of affect elevation specific to each task.

Associations between behavioral tasks and affect. Greater errors per second on the mirror tracing task (indicative of lower skill) was associated with shorter image persistence and shorter distress persistence on the EIT, providing initial evidence that the EIT and mirror tracing tasks are capturing related phenomena (see Table 3). However, mirror tracing total time was not associated with any of the EIT variables. In addition, shorter image persistence, shorter distress threshold and finding more of the slides distressing was associated with greater post-EIT negative affect even after controlling for baseline negative affect. Neither mirror tracing variable was significantly correlated with post-mirror tracing negative affect.

Self-report and behavioral tasks. None of the correlations between self-report and behavioral tasks were significant (see Table 3). It is also notable that discomfort intolerance was not correlated with the DTS, but was positively correlated with negative affect after the EIT.

Associations with psychopathology. We also measured depression and anxiety symptoms and binge eating symptoms. The mirror tracing variables did not correlate with either binge eating, or mood symptoms (see Table 3). Higher binge eating significantly correlated with shorter distress threshold. Although not significant, the magnitude of correlations between binge eating several of the other EIT variables (image persistence, count of distress) were similar to the distress threshold correlation. None

of the EIT variables were associated with mood symptoms. However, higher levels of anxiety and depression symptoms were correlated with lower self-reported distress tolerance.

Discussion

We expected that the EIT would demonstrate a small but significant association with the mirror tracing task, and that prediction was partially supported. The longer participants persisted on the EIT, the fewer errors per second they made on the mirror tracing task. It is notable that the skill variable was associated with the EIT whereas the total persistence time was not. It may be that errors per second better captures "frustration" or the emotional aspect of the mirror tracing task than simply how long a person was willing to persevere with the task.

We predicted that the EIT would show a stronger association with self-report measures of emotional distress tolerance than an existing persistence task, the mirror tracing task. Our predictions were not supported; neither the EIT nor the mirror tracing task correlated with the DTS. These results are consistent with other studies that have found no (or small) correlations between self-report and behavioral measures (Ameral et al., 2014; Bernstein et al., 2011; Cougle et al., 2013; Kiselica et al., 2015), even though the EIT was developed to be conceptually closer to the DTS. In other words, even though the EIT is in the same lower-order domain as the DTS, self-report and behavior do not converge, suggesting consistent discrepancies based on method of measurement.

Also, we found that the mirror tracing task skill index was subject to order effects but the EIT was not. One possible interpretation of this finding is that the EIT is less amenable to contextual influences than the mirror tracing task. The stability of EIT scores *should* be higher, as brief contextual influences should have less effect on an average of 45 slides than for tracing one difficult figure.

Finally, we have initial evidence that the EIT seems to demonstrate marginal associations with binge eating behaviors, where people with higher binge eating scores seemed to denote distress earlier into the images. This finding reflects prior research showing association between behavioral distress

tolerance measures and general impulsivity (Anestis et al., 2012), as binge eating is more associated with externalizing impulsive behaviors than internalizing behaviors.

Although there are certainly examples of the mirror tracing task used without incentives (Feldman et al., 2014; Macatee & Cougle, 2015), one limitation of the current study is that it may be difficult to generalize to other studies that used financial incentives for the mirror tracing task. Whether financial incentives increase ecological validity by making the task more similar to real-world instances of frustrating goal pursuit (e.g., completing a frustrating work assignment) is a question that has not yet been addressed empirically, and would be valuable for future research. Another limitation of the current study is the appearance of order effects. The order effects suggest that potentially having a greater space between the task presentations, such as by doing a neutral "filler" task, might allow participants sufficient time to recover before embarking on a second task.

Study 3

The function of Study 3 was to compare the EIT to a physical discomfort tolerance task. In addition, we wanted to expand the assessment of self-reported distress tolerance by including a newer measure that assesses tolerance to specific negative emotional states, more specific than simply tolerance to the generalized "distress" (Bernstein & Brantz, 2012).

Method

Participants. Similar recruitment procedures were used as the prior studies, where we determined eligibility from an online screener. For this study, participants were recruited from the subject pool and via flyers from the general undergraduate population, the latter during summer months when the subject pool was unavailable. In total, 1156 participants were screened, and 886 were eligible. There were no differences between recruitment groups on gender or age, but a greater percentage of the general undergraduate group was ineligible (30.03%) compared to the subject pool (20.77%), $\chi^2 = 11.16$, $p = .001$. We opened participation up for 100 sessions for the eligible participants, and 98 participated in the current study (58.2% female, 73.5% Caucasian, mean age = 19.67). Subject pool

participants (n = 53) were younger (M = 19.09, SD = 1.02) than the general undergraduate group (n = 45) M = 20.00, SD = 1.45), $t(96) = 3.57$, $p < .001$, but did not differ in gender, $\chi^2 = .01$, *ns.* Subject pool participants were compensated with course credit and general undergraduate participants were paid $10 for completing the study.

Measures.

Affect. For this study, we opted to measure emotion differently. Here, at baseline and after both tasks, participants were asked to report their mood during the task on a bipolar pleasantness scale ranging from 0 (*very unpleasant*) to 10 (*very pleasant*). After both tasks (but not at baseline), participants were also asked to report level of pain from 0 (*no pain*) to 10 (*a lot*) as a simple index of task pain severity.

Self-Reported Distress Tolerance. Similar to Study 2, we assessed self-reported distress tolerance with the DTS (Simons & Gaher, 2005), α = .91, and the Discomfort Intolerance Scale (DIS; Schmidt et al., 2006), α = .70. In addition, we used the Tolerance of Negative Affect States (TNAS; Bernstein & Brantz, 2012) a 24-item self-report measure assessing individual differences in perceived tolerance of specific negatively valenced affective states. TNAS items are rated on a 5-point Likert-type scale from 1 (*very intolerant*) to 5 (*very tolerant*), and subscales are grouped by specific emotion: Fear-Distress (α = .75), Sadness-Depression (α = .78), Anger (α = .89), Disgust (α = .85), Anxious-Apprehension (α = .50), and Negative Social Emotions (α = .73). Due to unacceptably low alpha, the Anxious-Apprehension scale was not used in analyses.

Psychopathology. Also similar to Study 2, we used the Binge Eating Scale (Gormally et al., 1982), α = .91, to assess symptoms of binge eating behavior, and the Depression, Anxiety, and Stress Scales (DASS; Henry & Crawford, 2005), α = .90, to assess mood symptoms.

Behavioral Distress Tolerance.

Cold pressor task. The cold pressor apparatus was a cooler half filled with water and ice. A metal grate separated the ice from the cold water, as direct skin-to-ice contact was not desired. Water

temperature was kept between 1 and 3 degrees Celsius (Quartana & Burns, 2007). Participants were asked to submerge their non-dominant hand up to the wrist in the water and to keep their hand still during the task. Participants were asked to verbally state when they first begin to feel pain. This initial pain assessment was the marker of cold pressor distress threshold, recorded in seconds. We also assessed how long (in seconds) a participant kept their hand in the water, up to a maximum of 300 seconds (5 minutes). A five minute limit was imposed on participants to reduce any physiological risk; the limit was not communicated to participants ahead of time (Burns, Bruehl, & Caceres, 2004).

Emotional image task. Same as in Study 2.

Procedure. After informed consent, participants completed baseline questionnaires via computer, followed by a pre-task assessment of mood via a one item bipolar pleasantness scale. Participants then completed both distress tolerance tasks in counterbalanced order, and then watched a funny video and were debriefed as in the previous studies.

Results

Data preparation. From the cold pressor tasks, three variables were calculated, to remain consistent with EIT scoring. Namely, we assessed the time until the participant removed his or her hand from the cold water, called cold pressor persistence. We also assessed discomfort threshold as the time until the participant noted pain or discomfort. Finally, we calculated the time between acknowledgment of pain and hand removal as distress persistence. Due to significant skewness and kurtosis in these indices, we used a log transformation for all three cold pressor variables (MacPherson et al., 2008). No order effects were found for cold pressor or EIT variables.

Associations among behavioral tasks. Higher scores on CP Persistence measures (total persistence and distress persistence) were significantly positively correlated with EIT persistence measures (see Table 3). However, the threshold measures across behavioral tasks were uncorrelated.

Affect change. Using a within subjects ANOVA, we compared baseline pleasantness and pleasantness after each task, controlling for order, $F(2, 186) = 68.94$, $p < .001$. We found that post-EIT

pleasantness was significantly lower (M = 2.52, SD = 2.12) than post-CP pleasantness (M = 3.95, SD = 2.26), which in turn was lower than baseline pleasantness (M = 6.84, SD = 1.78), all Bonferroni-corrected post-hoc tests significant at $p < .001$. All subsequent analyses involving pleasantness use post-task pleasantness scores controlling for baseline pleasantness to provide an indication of affect elevation specific to each task.

We also assessed differences in perceptions of pain after each task. We found no significant differences between physical pain after the cold pressor task (M = 7.53, SD = 1.94) and physical pain after the EIT (M = 7.49, SD = 1.76), $F(1, 96) = 2.50$, $p = .12$.

Associations between behavioral tasks and affect. In terms of association with post-task affect (see Table 4), longer EIT distress threshold and finding fewer slides distressing was associated with greater post-EIT pleasantness (after controlling for baseline pleasantness). Post-EIT pleasantness was not significantly associated with either EIT image persistence or EIT distress persistence. For the cold pressor task, longer overall persistence was associated with higher post-CP pleasantness, but CP distress threshold and CP distress persistence did not correlate with post-CP pleasantness. Higher pain was associated with lower persistence scores for both the EIT and the CP tasks.

Self-report measures. Also listed in Table 5 are the associations between self-report measures of distress tolerance and the behavioral tasks. Higher self-reported physical discomfort intolerance was associated with shorter CP persistence. In addition, higher self-reported disgust tolerance was associated with greater EIT image persistence and greater CP distress persistence. In addition, higher self-reported tolerance of sadness and anger was significantly associated with a longer EIT distress threshold. Of note, the general pattern of correlations between the TNAS anger, sadness and disgust subscales and the EIT variables is similar in terms of direction and magnitude, even though only a few of the correlations reached statistical significance. None of the EIT or CP variables correlated significantly with any of the symptom measures (depression, anxiety or binge eating).

Discussion

The results of Study 3 demonstrate concordance between the EIT and the cold pressor task; both overall persistence and distress persistence (e.g., persistence after distress was denoted) for each task were associated with the persistence measures on the other task. Both tasks were also associated with pain, such that higher persistence on the tasks was associated with lower post-task pain. These findings demonstrate further evidence for concurrent validity of the EIT.

However, we again found little relationship between self-report and behavioral measures of DT, particularly with the standard self-report measure (DIS, DTS, TNAS); neither the cold pressor task nor the EIT correlated with the DTS, although higher cold pressor persistence was associated with lower self-reported discomfort intolerance on the DIS. Compared to the CP, the EIT seemed to demonstrate slightly stronger relations to the TNAS scales, which measure tolerance to specific negative affective states. The EIT includes images associated with anger (e.g., World Trade Center images from September 2011), sadness (e.g., abandoned children), and disgust (e.g., dirty toilets and mutilated bodies), so it makes sense that these specific scales of the TNAS would be associated with the EIT. Indeed, the EIT does not really include images depicting fear or social emotions, so these results provide some evidence of content validity. Perhaps people who find anger, sadness and disgust easier to tolerate find those emotions easier to handle because they don't attend to them as quickly, as evidenced by slower identification of images as distressing. Alternately, people who quickly recognize distress in images, who may be more highly attuned to emotionally evocative stimuli (e.g., as has been found with borderline personality disorder symptoms; Domes, Schulze, & Herpertz, 2009) may also acknowledge more difficulties tolerating the common negative emotions of sadness, anger and disgust.

There is also some evidence that self-reported tolerance of disgust may be associated with persistence—this was the only index of the TNAS that was significantly associated with persistence scores (image persistence on the EIT and distress persistence on the cold pressor task). Even beyond statistical significance, the magnitudes of correlations between self-reported disgust tolerance and persistence measures were relatively consistent across the tasks—this was the only self-report measure

that even came close to significantly predicting EIT distress persistence (.20 was marginally significant). Why would tolerance to *disgust* in particular relate to persistence? Placing a hand in cold water doesn't seem "disgusting" in the traditional sense, and although the EIT includes images of disgust, it also includes images that are not inherently "disgusting." These results may reflect the function of disgust as helping people avoid things that might be dangerous (McNally, 2002). Alternatively, relationships with disgust may be particular to this sample. We note that the magnitude of the correlations was small and near the statistical significance cutpoint (e.g., correlations hovering around .20), suggesting small effects that may not be stable across samples; further replication of the function of self-reported tolerance to specific emotions (perhaps most specifically tolerance of disgust) in relation to behaviorally index distress tolerance is warranted.

General Discussion

Our central intention when developing the Emotional Image Tolerance (EIT) task was to develop a behavioral index of tolerance to negative emotion to expand measurement of this particular distress tolerance lower-order construct. In doing so, we examined the conceptual gaps between self-report and behavioral measures (e.g., breath of "distress" state, persistence versus experience, etc.) and tried to reduce them, with the goal of furthering both empirical and conceptual discussions about the operationalization of distress tolerance. In particular, we hoped to develop a task that assessed tolerance of a diffuse negative emotional state, not just narrow negative states of physical pain or frustration. We wanted a task with multiple trials which would ensure reliability and internal consistency. We also wanted to ensure that an actual experience of distress was present before being tolerated.

Were we successful in our endeavor? In some ways, yes. By selecting images from a well-validated image set such as the IAPS (Lang et al, 2008), that has been successfully used to induce negative affect for decades, the EIT addresses the lower-order domain of tolerance to negative emotion more explicitly than existing tasks. With 45 images in the set, we were also able to determine high internal consistency (Study 1). Moreover, our operationalization of DT ensured that we were able to

measure both overall persistence, an index consistent with extant persistence tasks, *and* tolerance of distress while persisting.

Associations Between EIT and Other Measures

Our prediction was that the EIT would exhibit small but significant correlations with existing behavioral tasks, providing evidence of validity of the EIT as a useful index of persistence. We *did* find significant relationships between EIT measures and both the mirror tracing skill index (Study 2) and the cold pressor (Study 3). Considering that all of these measures use an iteration of "time to giving up" as the index of tolerance, it is important that the EIT correlated with the extant tasks. At a very basic level, the persistence tasks assess the aspect of "stick-to-it-ness" within distress tolerance.

We also thought that due to the careful construction of the EIT to map onto -report measures (e.g., the DTS), that the EIT could demonstrate significant but low correlations with self-report measures. However, our findings here were inconsistent. Lower distress threshold was associated with higher physical discomfort intolerance (DIS) in Study 2[2] whereas greater image persistence was associated with higher self-reported distress tolerance (DTS) in Study 1. Other correlations which failed to reach statistical significance were around the same magnitude (.20 to .36) as the significant correlations, and in the predicted directions. Consistent with other work showing non-significant or tenuous associations between self-report and behavioral measures (Ameral et al., 2014; Bernstein et al., 2011; Cougle et al., 2013; Kiselica et al., 2015), we found lack of consistent relationships across measurement methods, even in assessments within the same lower-order domain. The major difference with the current set of studies is that the EIT was developed intentionally to illuminate differences

[2] A fourth study (n = 83) was conducted to evaluate the effect of an ego depletion manipulation on the EIT, which we have not included here due to concerns about the robustness of the depletion effect (Hagger & Chatsizarantis, 2015), exacerbated by findings that our depletion manipulation was ineffective (full description of the study available in Supplemental materials online). We also assessed symptoms of psychopathology in the baseline measures given at the start of Study 4 and found that higher binge eating was significantly (p < .05) associated with lower EIT image persistence (r = -.25), lower EIT distress threshold (r = -.30), and finding more images distressing (r = .29). Higher depressive symptoms were also associated with lower image persistence (r = -.25), and finding more slides distressing (r = .25). In addition, the DTS significantly correlated with both distress threshold (r = .20) and distress count (r = -.23). The DIS also significantly correlated with image persistence (r = -.30), distress threshold (r = -.36) and count of distress (r = .37).

between self-report and behavioral measures. The fact that we still did not find significant relationships across method suggests that method of assessment is a likely factor in non-concordance, rather than construct differences.

We also found associations between the EIT variables and symptoms of psychopathology, although the relationships we found were inconsistent and small in magnitude. Yet even in Study 2, when the correlations between depressive symptoms and EIT did not reach significance, they were greater in magnitude than associations between mirror tracing scores and psychopathology symptoms. Inconsistences in these correlations may be due to the fact that both Studies 2 and 3 had multiple tasks, because the samples excluded people with heightened current distress, or perhaps because the EIT does not consistently predict measures of psychopathology. Additional work on the EIT and psychopathology, perhaps with specific clinical samples (Bernstein et al., 2011; Kiselica et al., 2015), may be warranted. Because our symptom measures were self-report, obtaining indices of psychopathology from other methods (e.g., interviews, disorder-specific behaviors) may be particularly useful in future work with the EIT.

Tolerating Emotion

One unique aspect of the EIT, as compared to other behavioral measures of distress tolerance, is that the EIT included a novel assessment of distress persistence. In addition to using the average time the participant was willing to engage with the slide (e.g., image persistence), an index which is similar to the other behavioral task persistence time measures, the EIT also evaluated persistence *after* acknowledgment of distress. This allows the EIT to get closer to "persisting in the feeling of the emotion itself" while still focusing on persistence on an emotional task.

As should be expected, image persistence correlated strongly and positively with distress persistence; both are measures of persistence with the same images. Yet, distress persistence—which is conceptually the closest to the definition of distress tolerance—demonstrated far fewer relationships with included variables than the variable of image persistence. There are several possible reasons for

this finding. One is that the distress persistence index used in these studies was restricted to only slides in which the participant denoted distress, which means fewer slides went into the calculation and the analyses excluded any participants who simply watched all the slides go by (e.g., never indicating distress or a desire to move on to the next slide faster). However, considering that the entire EIT included 45 slides and most people indicated distress on a substantial number of images, this is not a particularly likely reason.

Second, as conceptualized here, the distress persistence index had a restricted range. The images could be viewed for up to 30 seconds, and on average distress was noted (i.e., distress threshold) at about 10 seconds into the slide. That leaves only 20 seconds for participants to tolerate the slide. An individual person who waited until 18 seconds to note an image as distressing then only had 12 seconds to tolerate the slide until the program automatically moved to the next image. It may be that such a person could have tolerated the image for longer, but the task did not allow this; the distress persistence measure is thus probably underestimating tolerance, particularly for people more capable of sitting with these negative images. Extending the ceiling for the task, or reconstructing the task to allow for 30 full seconds to *tolerate* the image after denoting distress would be possible in a revised iteration of the EIT, but we intentionally chose this method (despite its limitations) to standardize the amount of time each individual could look at each image, recognizing that some limit would have to be set to allow the task to continue even for people who do not find these images upsetting.

We also recognize that the distress persistence index, because it relies on distress threshold, also relies on awareness of emotion and willingness to report emotion. People vary in their awareness of their feelings; this is an aspect of emotional intelligence (Salovey & Grewal, 2005) and associated with emotion regulation (Boden & Thomspon, 2015). People with difficulties recognizing their feelings or people who are unwilling to report on their feelings almost certainly responded to the task differently than people aware of and willing to acknowledge emotion. Considering that self-reported emotion is the cornerstone of both basic emotion and emotion regulation research, the current set of studies is not

unique with regards to this limitation, but it is still worth noting, and potentially measuring emotional awareness in future work using the EIT.

Because distress persistence comes conceptually closest to the idea of tolerating emotional distress (e.g., sitting with distress that is actually present), we actually find the *lack* of significant relationships quite interesting, and worthy of reflection. If we assume that the EIT distress persistence index assesses something meaningful and conceptually closer to what is assessed by self-report scales, why does it not correlate with self-reported distress tolerance? We raise this question here with recognition that answering it will require examination of the lingering conceptual distinctions between self-reported and behavioral indices of distress tolerance, which was one of the central goals of the EIT.

Lingering Conceptual Distinctions

First, all of the behavioral tasks—including the EIT—are measures of persistence. They assess how long someone continues to engage with a task. Using persistence tasks makes sense because that is essentially what the word tolerance implies—willingness to "put up with" something without avoidance or escape. Yet, many of the self-report measures of distress tolerance are measures of distress *intolerance,* with items like "I can't bear disturbing feelings" on the Frustration Discomfort Scale (Harrington, 2005) or "There's nothing worse than feeling distressed or upset" on the DTS (Simons & Gaher, 2005). These items do not address persistence, but rather assess judgments about feelings. Only a few items on the extant self-report measures actually address the kind of situations that are tapped by behavioral tasks. For one example, the Discomfort Intolerance Scale has the item "When I begin to feel physically uncomfortable, I quickly take steps to relieve discomfort" (Schmidt et al., 2006). Similarly, items on the regulation subscale of the DTS get at efforts to alleviate discomfort, though not as directly: "I'll do anything to avoid feeling distressed or upset" (Simons & Gaher, 2005). Essentially, although tolerance and intolerance are conceptual opposites, neither are synonymous with persistence. Judgments about emotions are also not the same as willingness to continue engaging with distress, although clearly these concepts overlap. This may be why the EIT does not correlate with existing self-report measures.

Prior work has indicated that the existing behavioral tasks are more narrow than self-report measures (Kiselica et al., 2015) but have not clearly described how the behavioral tasks differ in scope. Our view is that the behavioral tasks are narrowly focused on persisting through a task that causes the distress, which is qualitatively different than sitting with free-floating distress or distress caused by another source. For example, if a woman has a fight with her romantic partner, she might feel angry and/or sad. Yes, perhaps one of her goals might be to persist in the discussion with her partner without storming out or saying something that she would regret later (Linehan, 1993). She might also need to find a way to tolerate her feelings that linger after the fight itself is over, such as when she recalls elements of the argument, or when she has negative appraisals about her partner or about herself. Another goal might be persisting through *other* tasks after the fight. If the woman has a presentation at work the day of her fight, she might have to tolerate her feelings while still accomplishing her work. Provided persistence is important to the study of distress tolerance, what must be persisted? The situation that evoked the emotion, the feeling of the emotion itself, or other tasks that might be more difficult in the face of a strong emotion evoked elsewhere? Self-report measures of distress tolerance seem to focus on tolerating the feeling, whereas the behavioral measures of distress tolerance seem to focus on tolerating the emotion-generating task. It is here where the EIT is particularly unique, because with the EIT the task *itself* is clearly emotional, and the distress persistence index at least conceptually captures that idea of tolerating the feeling of distress, not just the task itself.

Finally, assessing behavior means assessing an action at a given moment in time. Self-report measures typically ask for aggregated summaries across time and situations, but behavior—by necessity—is assessed at one moment in time and in one specific context. This is consistent with the often-found "state" versus "trait" contrast in assessment, whereby self-report measures typically get at more "trait-like" tendencies, and behavioral tasks typically assess more "state-like" momentary abilities (Mischel, 1968; Podsakoff et al., 2003). In addition, within the overall area of distress tolerance, each task assesses specific behaviors, and can be viewed as assessing one lower-order domain (Zvolensky et

al., 2010). Indeed, the significant associations between behavioral measures found in the extant literature are typically between measures in the same domain, such as frustration tolerance, where the PASAT and mirror tracing typically correlate significantly (Ameral et al., 2014; Bornovolova et al, 2008; Daughters et al., 2005; Kiselica et al., 2015; MacPherson et al., 2008; McHugh et al., 2011). Yet, correlations of behaviors *between* domains (e.g., frustration tasks with physical tasks) are often non-significant (Brown et al., 2002; Glassman et al., 2016; MacPherson et al., 2008; Szuhany & Otto, 2015). Some authors have suggested that behavioral tasks may be most helpful in identifying people with inflexible and generalized distress intolerance (Ameral et al., 2014). We certainly agree that people who have difficulties tolerating distress across a variety of situations and contexts are likely targets of treatment. We also think that variability in behaviorally indexed distress tolerance may be a useful direction for future research, recognizing that variability is not inherently the same as *flexibility* (Aldao, Sheppes & Gross, 2015). In addition, people who are not always intolerant of distress may find themselves with difficulties sitting with distress on occasion, such as when life stressors are particularly high or when self-care (e.g., sleep, nutrition) is low; understanding the factors that contribute to momentary distress tolerance for these people is also a valuable research endeavor. Ultimately, even the EIT, which was designed to be more reliable and thus less influenced by context, is still one specific behavior (e.g., tolerating negative emotional images) given at one laboratory session and situated at one moment in time. Capitalizing on behavioral tasks and the unique benefits that they bring—even if they do *not* overlap with self-report measures—may be a useful direction for future research.

Limitations

There are several limitations to our current work. First, we recognize that by restricting our sample to those not experiencing as much current distress, we are likely downplaying the relationships amongst the distress tolerance measures and with symptom variables. Including a sample with more current distress, potentially a sample with greater severity in pathology history (e.g., history of suicide attempt or non-suicidal self-injury, history of drug use, symptoms of borderline personality disorder)

will be important in future work. Notably, although we might expect a clinical sample to respond more intensely to the EIT (e.g., more slides found distressing, lower distress threshold, lower persistence), we would not expect qualitatively different responses, although this remains an empirical question.

Second, we recognize that with a visual image task, ensuring that participants look at the images is essential. Although our experimenters watched participants to prevent them from actively turning away, more sophisticated methods such as eye tracking could be used to assess image attention. This would be particularly useful in examining where participants direct their attention *within* the image, considering research showing that looking at non-arousing areas of negative images decreases emotional responsivity (Dunning & Hajcak, 2008; Van Reekum et al., 2007).

We also recognize that there are a myriad of choices when developing a new task, both in creating the task itself and in the measures used to assess validity. In terms of the task itself, we chose to use a persistence-based measure because the other behavioral tasks use persistence measures. But is persistence the best index of tolerance? People persist or fail to persist on distress tolerance tasks for all kinds of reasons, only some of which are related to distress (Ameral et al., 2014), and are likely associated with other emotional goals (Tamir, 2016). We also recognize that we chose somewhat unorthodox measures for psychopathology (e.g., depression, anxiety, binge eating) rather than those which have shown greater relationship to distress tolerance in the past (e.g., borderline symptoms, drug history). We chose these based on our sample and because we thought the EIT might be associated with them, though in retrospect we recognize that other additional measures might have been stronger choices.

Conclusion & Future Directions

Our impetus for developing the EIT came from a recognition that a behavioral task assessing tolerance to negative emotion might fill a gap in the assessment one lower-order domains within the distress tolerance construct (Zvolensky et al., 2010). We also wanted to better understand how method invariance and construct distinctness contribute to the often-found lack of association between self-

report and behavioral measures by designing a measure intentionally to bridge the gap. We recognize that people's self-perceptions of themselves are unlikely to ever map exactly onto observable behaviors, and method invariance is likely to continue. Yet, the process of attempting to understand the nuanced ways in which behavioral and self-report measures of distress tolerance differ provides several avenues for future research. Our work provides a pathway for future researchers to investigate tolerance to negative emotion from a behavioral perspective, whether via using the EIT or by using the ideas advanced here to develop other behavioral measures of tolerance to negative emotion, such as measures explicitly assessing tolerance to individual negative emotional stages (e.g., anger, sadness), or tolerance to idiographic emotional stimuli (e.g., images or objects that hold personal significance for an individual). It also may be important to go the other direction, by developing self-report measures that better map onto the goal-oriented, persistence-based behavioral measures. These new self-report measures would likely require current, momentary assessment of distress tolerance rather than retrospective reports (Cougle et al., 2013).

This work may also have significant clinical utility, with fruitful potential future research directions. That is, do scores on this task predict important clinical outcomes, such as the usefulness of learning distress tolerance skills during psychotherapy (Linehan, 1993), or anticipating difficulties that might be encountered during exposures? This task could also be modified as a clinical tool for *increasing* distress tolerance, such as having a client engage in this task during a psychotherapy session, with a therapist coaching a client how to better tolerate and thus persist in looking at the disturbing images. Further investigation of the EIT in the clinical arena, both in terms of how the EIT corresponds with psychopathology and how the EIT may be explicitly used as a clinical tool, is warranted.

Ultimately we concur with others who have stated that distress tolerance is an important transdiagnostic risk factor for psychopathology (Leyro et al., 2010; Zvolensky et al., 2010). To fully understand how the specific lower-order domains of distress tolerance differ from one another, and may differentially predict outcomes indicative of both pathology and well-being, additional attention to

measurement can assist in the goal of conceptual refinement. By doing this kind of work, we can move toward a clearer understanding of distress tolerance and it's components, which will then allow for a more nuanced understanding of distress tolerance as a transdiagnostic risk factor.

References

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

Ameral, V., Reed, K. M., Cameron, A., & Armstrong, C. (2014). What are measures of distress tolerance really capturing? A mixed methods analysis. *Psychology of Consciousness, 1,* 357-369. doi: 10.1037/cns0000024

Anestis, M. D., Lavender, J. M., Marshall-Berenz, E. C., Gratz, K. L., Tull, M. T., & Joiner, T. E. (2012). Evaluating distress tolerance measures: Interrelations and associations with impulsive behaviors. *CognitiveTherapy and Research*, 36(6), 593-602. doi: 10.1007/s10608-011-9377-8

Bem, D. J. (1972). Self-perception theory. *Advances in Experimental Social Psychology, 6,* 1-62. doi: 10.1016/S0065-2601(08)60024-6

Bernstein, A., & Brantz, H. (2013). Tolerance of Negative Affective States (TNAS): Development and evaluation of a novel construct and measure. *Cognitive Therapy and Research, 37,* 421-433. doi:10.1007/s10608-012-9471-6

Bernstein, A., Marshall, E. C., & Zvolensky, M. J. (2011). Multi-method evaluation of distress tolerance measures and construct(s): Concurrent relations to mood and anxiety psychopathology and quality of life. *Journal of Experimental Psychopathology, 2,* 386-399. doi: http://dx.doi.org/10.5127/jep.006610

Boden, M. T., & Thompson, R. J. (2015). Facets of emotional awareness and associations with emotion regulation and depression. *Emotion, 15*(3), 399-410. doi: 10.1037/emo0000057

Bornovalova, M., Gratz, K., Daughters, S., Nick, B., Delany-Brumsey, A., & Lynch, T. (2008). A multimodal assessment of the relationship between emotion dysregulation and borderline

personality disorder among inner-city substance users in residential treatment. *Journal of Psychiatric Research, 42* (9) (2008), pp. 717–726. doi: 10.1016/j.jpsychires.2007.07.014

Brown, R. A., Lejuez, C. W., Kahler, C. W., & Strong, D. R. (2002). Distress tolerance and duration of past smoking cessation attempts. *Journal of Abnormal Psychology, 111,* 180-185. doi: 0.1037//0021-843X.111.1.180

Burns, J. W., Bruehl, S., & Caceres, C. (2004). Anger management style, blood pressure reactivity, and acute pain sensitivity: evidence for "trait× situation" models. *Annals of Behavioral Medicine, 27*(3), 195-204. doi: 10.1207/s15324796abm2703_7

Coan, J. A. & Allen, J. J. (Eds) (2007). *Handbook of Emotion Elicitation and Assessment.* New York: Oxford University Press.

Cochran, C. D., & Hale, D. (1985). College student norms on the brief symptom inventory. *Journal of Clinical Psychology, 41,* 777-779.

Cougle, J. R., Bernstein, A., Zvolensky, M. J., Vujanovic, A. A., & Macatee, R. J. (2013). Validation of self-report measures of emotional and physical distress tolerance. *Journal of Psychopathology and Behavioral Assessment, 35,* 76–84. doi:10.1007/s10862-012-9317-2

Daughters, S. B., Lejuez, C., Bornovalova, M. A.,Kahler, C. W., Strong, D. R., & Brown, R. A. (2005). Distress tolerance as a predictor of early treatment dropout in a residential substance abuse treatment facility. *Journal of Abnormal Psychology,114,* 729–734. doi:10.1037/0021-843X.114.4.729

Daughters, S. B., Sargeant, M. N., Bornovalova, M. A., Gratz, K. L., & Lejuez, C. W. (2008). The relationship between distress tolerance and antisocial personality disorder among male inner-city treatment-seeking substance users. *Journal of Personality Disorders, 22,* 509–524. doi:10.1521/pedi.2008.22.5.509

Derogatis, L. R., & Melisaratos, N. (1983). The brief symptom inventory: An introductory report. *Psychological Medicine, 13*(3), 595-605.

Domes, G., Schulze, L., & Herpertz, S. C. (2009). Emotion recognition in borderline personality disorder—A review of the literature. *Journal of Personality Disorders, 23,* 6 – 19.

Dunning, J. P., & Hajcak, G. (2008). See no evil: Directing visual attention within unpleasant images modulates the electrocortical response. *Psychophysiology, 46,* 28-33. doi: 10.1111/j.1469-8986.2008.00723.x

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3,* 364-370. doi: 10.1177/1754073911410740.

Ellard, K. K., Farchione, T. J., & Barlow, D. H. (2012). Relative effectiveness of emotion induction procedures and the role of personal relevance in a clinical sample: A comparison of film, images, and music. *Journal of Psychopathology and Behavioral Assessment, 34,* 232-243. doi: 10.1007/s10862-011-9271-4

Feldman, G., Dunn, E., Stemke, C., Bell, K., & Greeson, J. (2014). Mindfulness and rumination as predictors of persistence with a distress tolerance task. *Personality and Individual Differences, 56,* 154-158. doi: 10.1016/j.paid.2013.08.040.

Fleeson, W. (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*(6), 1011-1027. doi: 1037/0022-3514.80.6.1011

Glassman, L. H., Martin, L. M., Bradley, L. E., Ibrahim, A., Goldstein, S. P.., . . . & Herbert, J. D. (2016). A brief report on the assessment of distress tolerance: Are we measuring the same construct? *Journal of Rational-Emotive and Cognitive-Behavior Therapy, 34,* 87-99. doi:10.1007/s10942-015-0224-9

Gormally, J., Black S., Daston S., & Rardin D., C.D. (1982). The assessment of binge eating severity among obese persons. *Addictive Behaviors, 7*(1), 47-55.

Hagger, M. S.,& Chatzisarantis, N. L. (2015). A multi-lab pre-registered replication of the ego-depletion effect. *Perspectives on Psychological Science, 11*, 546-573. doi: 10.1177/1745691616652873

Harrington, N. (2005). The Frustration Discomfort Scale: Development and psychometric properties. *Clinical Psychology & Psychotherapy, 12*(5), 374-387.

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21: Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227-239. doi: 10.1348/014466505X29657

Kiselica, A. M., Rojas, E., Bornovalova, M. A., & Dube, C. (2015). The nomological network of self-reported distress tolerance. *Assessment, 22*, 715-729. doi: 10.1177/1073191114559407

Kiselica, A. M., Webber, T., & Bornovalova, M. A. (2014). Stability and change in distress tolerance and it's prospective relationship with borderline personality features: A short-term longitudinal study. *Personality Disorders: Theory, Research, & Treatment, 3*, 247-256. doi: 10.1037/per0000076

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International Affective Picture System (IAPS): *Affective ratings of pictures and instruction manual.* Technical Report A-8. Gainesville, FL: University of Florida.

Lejuez, C.W., Kahler, C.W., & Brown, R. A. (2003). A modified computer version of the Paced Auditory Serial Addition Task (PASAT) as a laboratory-based stressor. The *Behavior Therapist, 26*(4), 290−293.

Leyro, T. M., Zvolensky, M. J., & Bernstein, A. (2010). Distress tolerance and psychopathological symptoms and disorders: a review of the empirical literature among adults. *Psychological Bulletin, 136*(4), 576–600. doi:10.1037/a0019712

Linehan, M. (1993). *Cognitive-behavioral Treatment of Borderline Personality Disorder.* Guilford press: New York.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*(3), 335-343.

Macatee, R. J., & Cougle, J. R. (2015). Development and evaluation of a computerized intervention for low distress tolerance and its effect on performance on a neutralization task. *Journal of Behavior Therapy and Experimental Psychiatry, 48,* 33-39. doi: 10.1016/j.jbtep.2015.01.007

MacPherson, L., Stipelman, B. A., Duplinsky, M., Brown, R. A., & Lejuez, C. W. (2008). Distress tolerance and pre-smoking treatment attrition: Examination of moderating relationships. *Addictive Behaviors, 33,* 1385-1393. doi: 10.1016/j.addbeh.2008.07.001

McHugh, R. K., Daughters, S. B., Lejuez, C. W., Murray, H. W., Hearon, B. A., Gorka, S. M., & Otto, M. W. (2011). Shared variance among self-report and behavioral measures of distress intolerance. *Cognitive Therapy and Research, 35*(3), 266–275. doi:10.1007/s10608-010-9295-1.

McHugh, R. K., & Otto, M. W. (2011). Doman-general and domain-specific strategies for the assessment of distress intolerance. *Psychology of Addictive Behaviors, 25,* 745-749. doi: 10.1037/a0025094

McNally, R. J. (2002). Disgust has arrived. *Anxiety Disorders, 16,* 561-566.

Mischel, W. (1968). *Personality and assessment.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259. http://dx.doi.org/10.1037/0033-295X.84.3.231

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903. doi: 10.1037/0021-9010.88.5.879

Quartana, P. J., & Burns, J. W. (2007). Painful consequences of anger suppression. *Emotion, 7*(2), 400 - 414. Doi: 10.1037/1528-3542.7.2.400.

Ross, L., & Nisbett, R.E. (1911). *The person and the situation.* Philadelphia, PA: Temple University Press.

Salovey, P., & Greweal, D. (2005). The science of emotional intelligence. *Current Directions in Psychological Science, 14,* 281-285.

Schmidt, N.B., Richey J.A., & Fitzpatrick, K.K. (2006). Discomfort intolerance: Development of a construct and measure relevant to panic disorder. *Journal of Anxiety Disorders, 20(*3), 263-280. doi: 10.1016/j.janxdis.2005/02.2002

Simons, J. S., & Gaher, R. M. (2005). The Distress Tolerance Scale: Development and validation of a self-report measure. *Motivation and Emotion, 29,* 83-102. doi: 10.1007/s11031-005-7955-3

Strong, D., Lejuez, C., Daughters, S., Marinello, M., Kahler, C., & Brown, R. (2003). The computerized mirror tracing task. [Unpublished manual]. Available from http://www.addiction.umd.edu/downloads.htm.

Szuhany, K. L., & Otto, M. W. (2015). Contextual influences on distress intolerance: Priming effects on behavioral persistence. *Cognitive Therapy & Research, 39,* 499-507. doi: 10.1007/s10608-015-9672-x

Tamir, M. (2016). Why do people regulate their emotions? A taxonomy of motives in emotion regulation. *Personality and Social Psychology Review, 20,* 199-222. doi: 0.1177/1088868315586325

Van Reekum, C. M., Johnstone, T., Urry, H. L., Thurow, M. E., Schaefer, H. S., Alexander, A. L., et al. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. *NeuroImage, 36,* 1041–1055.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070. doi: 10.1037/0022-3514.54.6.1063

Zvolensky, M. J., Vujanovic, A. A., Bernstein, A., & Leyro, T. (2010). Distress tolerance theory, measurement, and relations to psychopathology. *Current Directions in Psychological Science, 19*(6), 406-410.

Zvolensky, M. J.,Feldner, M. T., Eifert, G. H., & Brown, R. A. (2001). Affective style among smokers: Understanding anxiety sensitivity, emotional reactivity, and distress tolerance using biological challenge. *Addictive Behaviors, 26,* 901-915.

Table 1. *Image statistics from Study 1*

IAPS Slide Number	Percent Indicated Distress	Percent Viewed Entire Time	Average Distress Threshold	Average Distress Persistence	Average Total Slide Viewing Time	Mean Distress Rating
2095	43.48	30.43	7.14	9.47	16.61	2.79
2352	56.52	26.09	7.02	7.99	15.01	3.26
2375*	15.22	52.17	10.67	13.17	23.83	2.06
2800	39.13	41.30	8.08	9.87	17.94	2.77
3000	67.39	15.22	5.25	5.87	11.12	3.83
3001	52.17	26.09	6.13	7.99	14.12	3.51
3005	58.70	19.57	5.14	7.34	12.49	4.02
3015	54.35	30.43	6.98	8.51	15.49	3.40
3051	52.17	26.09	6.26	8.18	14.44	3.13
3053	63.04	15.22	4.72	7.04	11.76	4.00
3063	65.22	21.74	5.91	6.61	12.52	3.91
3064	45.65	26.09	6.74	9.01	15.75	3.30
3080	58.70	23.91	6.32	7.61	13.93	3.45
3100	50.00	19.57	5.16	8.26	13.42	3.55
3102	60.87	19.57	5.25	7.06	12.31	3.53
3120	47.83	32.61	7.52	8.96	16.48	2.87
3140	56.52	19.57	6.01	7.97	13.98	3.28
3150	60.87	19.57	6.17	7.46	13.63	3.68
3170	54.35	17.39	4.81	7.96	12.77	3.57
3181*	32.61	52.17	9.47	11.10	20.57	2.30
3191	41.30	34.78	7.22	9.39	16.61	2.79
3195*	41.30	39.13	8.54	10.44	18.99	2.36
3230*	23.91	58.70	10.22	12.38	22.60	2.15
3261	63.04	21.74	5.53	7.21	12.74	3.57
3266	60.87	19.57	5.23	6.87	12.10	3.57
3301	34.78	34.78	7.47	10.42	17.89	2.87
3350*	36.96	43.48	8.26	10.41	18.67	2.32
3530	28.26	45.65	8.89	11.67	20.56	3.00
6021	30.43	47.83	9.55	11.17	20.72	2.60
6022	41.30	28.26	7.02	9.48	16.51	2.66
6350	30.43	41.30	9.22	11.36	20.58	2.74
6360*	19.57	58.70	10.55	12.74	23.29	2.26
6415	43.48	36.96	7.96	9.07	17.03	2.77
6540	36.96	47.83	9.11	10.97	20.08	2.40
6560	32.61	52.17	9.64	11.03	20.68	2.68
6563	30.43	45.65	9.12	10.94	20.06	2.85
9040	41.30	32.61	7.20	10.06	17.26	2.68
9075	34.78	32.61	8.31	10.95	19.25	2.98

9140*	32.61	54.35	9.45	10.91	20.35	2.15
9163*	30.43	50.00	10.05	11.63	21.68	2.21
9181	39.13	30.43	7.43	9.99	17.42	2.64
9183	50.00	15.22	5.48	9.22	14.69	3.23
9187	50.00	21.74	6.66	8.84	15.50	3.36
9254*	28.26	52.17	9.69	11.55	21.24	2.40
9300*	52.17	36.96	7.52	8.32	15.84	2.32
9301	56.52	28.26	6.42	7.63	14.05	2.49
9322*	47.83	41.30	8.32	8.64	16.96	2.06
9325	50.00	32.61	6.51	7.96	14.47	2.62
9326*	45.65	45.65	7.99	8.79	16.78	2.21
9405	47.83	28.26	6.54	8.62	15.16	3.30
9410	56.52	17.39	5.77	8.47	14.24	3.55
9412	39.13	45.65	9.55	9.97	19.52	2.47
9413*	28.26	47.83	9.67	11.65	21.32	2.36
9433	36.96	39.13	8.35	10.12	18.47	2.55
9435*	19.57	67.39	11.69	12.68	24.37	2.02
9570	41.30	45.65	8.39	9.41	17.80	2.49
9635	47.83	36.96	7.60	9.18	16.78	3.26
9921	34.78	36.96	8.16	10.90	19.06	2.91
9940	30.43	39.13	8.48	11.04	19.53	3.00

**not retained in final version of the task*

Table 2. *Zero order correlations between all Study 1 variables*

		1.	2.	3.	4.	5.	Mean (SD)
1.	EIT Avg. Distress Threshold	--					7.05 (4.36)
2.	EIT Distress Persistence	.76**	--				9.00 (5.35)
3.	Self-Report Distress Tolerance (DTS)	.28	.34*	--			3.88 (.79)
4.	Post-task NA[a]	-.33*	-.27	-.21	.--		16.08 (4.50)
5.	Post-task PA[a]	.21	.06	-.16	-.32*		21.08 (7.53)
6.	Distress Ratings	-.66**	-.60**	-.11	.61**	-.46**	3.11 (1.67)

*$p < .05$, **$p < .01$

EIT = Emotional Image Tolerance; DTS = Distress Tolerance Scale; NA = PANAS Negative Affect; PA = PANAS Positive Affect

[a]Correlations involving post-task affect are partial correlations controlling for baseline affect

Table 3. *Correlations between all Study 2 variables*

		1.	2.	3.	4.	5. [a]	6.	7. [b]	8. [b]	9.	10.	11.	12.
1.	EIT Avg Image Persistence	--											
2.	EIT Avg Distress Threshold	.86**	--										
3.	EIT Count Distress	-.79**	-.97*	--									
4.	EIT Avg Distress Persistence	.76**	.42**	-.38**	.--								
5.	MT Errors per Second (log transformed)[a]	-.29*	-.22	.20	-.24*	.--							
6.	MT Total Time (log transformed)	.12	.06	-.10	.11	-.09							
7.	Post-EIT Negative Affect[b]	-.27*	-.35**	.34**	-.05	.16	-.02	.--					
8.	Post-MT Negative Affect[b]	-.15	-.23	.26*	.01	.06	.08	.43**	--				
9.	Self-Report Distress Tolerance (DTS)	-.02	.06	-.07	-.10	.04	.22	-.12	-.13	--			
10.	Self-Report Discomfort Intolerance (DIS)	-.22	-.22	.17	-.16	.14	.14	.25*	.22	-.16	--		
11.	Binge Eating (BES)	-.21	-.22*	.22	-.06	.11	-.04	.23*	.18	-.18	.02	--	
12.	Depression (DASS)	-.13	-.17	.16	.02	-.02	.11	.43**	.24	-.47**	.19	.20	--
13.	Anxiety (DASS)	-.09	-.16	.19	.02	-.05	-.11	.33**	.19	-.54**	.12	.23	.65**

*$p < .05$, **$p < .01$

EIT = Emotional Image Tolerance; MT=Mirror Tracing Persistence Task; NA = PANAS Negative Affect; DTS = Distress Tolerance Scale; DIS = Discomfort Intolerance Scale; BES = Binge Eating Scale; DASS = Depression, Anxiety & Stress Scales

Table 4. *Correlations between EIT task and CP task variables in Study 3*

	EIT Avg. Image Persistence	EIT Avg Distress Threshold	EIT Count Distress	EIT Avg Distress Persistence
CP Persistence	.30**	.21*	-.14	.27**
CP Distress Threshold	.07	.09	-.10	.03
CP Distress Tolerance	.26*	.11	-.06	.32*

*$p < .05$, **$p < .01$

EIT = Emotional Image Tolerance; CP = Cold pressor

Table 5. *Correlations between Study 3 distress tolerance tasks, state variables and self-reported distress tolerance measures*

	EIT Avg. Image Persistence	EIT Avg Distress Threshold	EIT Count Distress	EIT Avg Distress Persistence	CP Persistence[c]	CP Distress Threshold[c]	CP Distress Tolerance[c]
Post-Task Pleasantness[ab]	.18	.35*	-.36*	-.09	.22*	.10	.18
Post-Task Pain[a]	-.28*	-.19	.16	-.26*	-.30**	-.02	-.18
Self-Report Distress Tolerance (DTS)	.10	.11	-.10	.04	-.01	-.16	.08
Self-Report Discomfort Intolerance (DIS)	.01	.02	.06	.01	-.23*	-.05	-.19
TNAS-Fear-Distress	.07	.10	-.12	-.09	.07	-.09	.13
TNAS-Sadness	.18	.22*	-.22*	.03	.10	-.12	.11
TNAS-Anger	.13	.21*	-.17	-.04	.09	-.02	.04
TNAS-Disgust	.21*	.17	-.16	.20	.14	.08	.21*
TNAS-Negative Social Emotions	.004	.06	-.05	-.09	.04	-.04	.07
Binge Eating	-.09	-.18	.18	.000	-.07	.004	.008
Depression	.11	.14	-.11	.09	.19	.02	.10
Anxiety	.01	-.03	.002	.07	.05	.10	-.07

*$p < .05$, **$p < .01$

EIT = Emotional Image Tolerance; CP = cold pressor task; DTS = Distress Tolerance Scale; DIS = Discomfort Intolerance Scale; TNAS = Tolerance of Negative Affective State

[a]Post-EIT state correlations are given for the EIT variables, post-CP state correlations given for the CP variables

[b]Correlations involving post-task pleasantness are partial correlations controlling for baseline pleasantness

[c]Variables are log-transformed.

Study 4 (Supplementary Material for Online)

Our intention when developing the EIT was to create a measure that was more trait-like, and thus less responsive to situational manipulations. To test this idea, we evaluated the effect of an ego depletion manipulation on the EIT. Ego depletion is the idea that using self-control subsequently leaves less self-control available for subsequent tasks (Muraven & Baumeister, 2000). The classic paradigm involves experimental condition participants engaging in a task that requires considerable effort (e.g., suppressing emotion or thoughts, attending to a target amidst distracting stimuli) while control participants engage in a related but less taxing task. Scores of studies have found that depletion participants perform worse on a second self-control task than control participants, indicating that engaging in self-control depletes self-control resources (Hagger, Wood, Stiff, & Chatzisarantis, 2010). Importantly, depletion studies often use persistence tasks such as the cold pressor or computerized persistence tasks as the secondary index of self-control (Hagger et al., 2010).

Thus, the current study used the EIT as the dependent variable in a traditional ego-depletion paradigm. If the EIT is more trait-like, as we expected, scores on the EIT should not be different after completing a depletion task than after completion of a control task.

Method

Participants. The same screening criteria were used, with 17 participants from the psychology subject pool who received course credit, and 66 college students who were paid $10. In total, 174 participants were screened, and 125 were eligible. The eligible participants were invited to complete a laboratory session, and of these, 83 participated in the current study (63.9% female, 62.7% Caucasian, mean age = 19.79). There were no differences in paid versus subject pool participants in age or gender, but a higher percentage of subject pool participants were minorities (n = 10, 59%) than paid participants (n = 21, 32%).

Procedure and Measures. After informed consent, participants completed baseline questionnaires via computer, including the Distress Tolerance Scale (DTS; Simons & Gaher, 2005, α

= .88), the Discomfort Intolerance Scale (Schmidt et al., 2006, α = .76), the Binge Eating Scale (BES; Gormally et al., 1982, α = .86), and the DASS-21 (Henry & Crawford, 2005, α = .80). These measures were followed by a baseline assessment of mood via the PANAS, where again we focus only on the NA scale (α =.78)

Participants were then randomized to depletion condition using the "crossing out e's" task (Baumeister et al., 1998; Tice et al., 2007). All participants were given a photocopied sheet from a statistics textbook and asked to cross out all instances of the letter 'e' on the page. This establishes the habit. Then, control participants were given a second page in which to cross out 'e's. Depletion participants were asked to complete the second page by crossing out all 'e's except 'e's that were one or two character spaces away from another vowel. Thus, the depletion condition participants have to exert greater self-control to override the learned responses from the first portion of the task. After the depletion task participants completed another PANAS (NA α =.66), followed by the EIT task (same as in Studies 2 and 3) and a final PANAS (NA α =.86). The end of the study involved watching a funny video and debriefing, just as in the prior studies.

Results

Our central question was whether EIT variables would change under depletion. There were no differences in any of the EIT variables based on depletion condition (all *ts* < 1). However, one possibility could be that depletion had a short-lived effect; it could have influenced initial EIT performance over the first few trials. All images were presented in 5 blocks of 9 slides each, which allowed us to calculate EIT variables by block. To assess the effects of depletion at different time points in the EIT, we conducted a mixed model ANOVAs with first and last block as the within subject factor and depletion condition as the between subjects factor, which also allowed us to assess any changes in EIT performance across the timeframe of the task. We found a main effect of block for image persistence, $F(1,79) = 13.07$, $p = .001$, where the persistence at block 5 ($M = 16.72$, $SD = 9.80$) was significantly shorter than at Block 1 ($M = 18.97$, $SD = 8.90$). A similar, albeit weaker,

main effect of block was evident for distress persistence, $F(1,74) = 5.66$, $p = .02$, where the persistence at block 5 ($M = 8.72$, $SD = 7.65$) was significantly shorter than at Block 1 ($M = 9.96$, $SD = 7.30$). Distress threshold also diminished from Block 1 ($M = 11.79$, $SD = 8.99$) to Block 5 ($M = 10.57$, $SD = 9.31$), $F(1,79) = 4.18$, $p = .04$. There was no change in distress count based on block ($F < 1$). Finally, there was no effect of condition or any condition by block interactions on any EIT variables. These results indicate that although participants became less persistent on the EIT over time, these shifts were not altered by depletion condition.

To ensure that EIT scores are relatively consistent across the task, we also performed correlations between Block 1 and Block 5 scores. Image persistence ($r = .83$, $p < .001$), distress threshold ($r = .83$, $p < .001$), count of distressing images ($r = .85$, $p < .001$), and distress persistence ($r = .82$, $p < .001$) were all strongly correlated, suggesting that even with a significant decline over time, the scores were still relatively consistent across the task.

We also found that positive affect changed over time, $F(2, 160) = 41.30$, $p < .001$, with a decrease from baseline ($M = 27.32$, $SD = 7.97$) to post-depletion ($M = 22.82$, $SD = 8.53$), and another decline in PA from post-depletion to post-EIT ($M = 20.47$, $SD = 8.29$) Negative affect did not significantly change from baseline ($M = 12.80$, $SD = 3.38$) to post-depletion ($M = 13.02$, $SD = 3.15$), but increased from post-depletion to post-EIT ($M = 17.29$, $SD = 6.16$), $F(2, 160) = 50.17$, $p < .001$. These changes in affect over time did not differ by depletion condition.

Change scores were calculated to index the change in affect from pre- to post-EIT. Correlations among all study variables are listed in Table 6. Higher discomfort intolerance and greater binge eating symptoms were both associated with shorter image persistence, shorter distress threshold and finding more images distressing. In addition, higher depressive symptoms were significantly or marginally associated with all of the EIT variables.

Discussion

The depletion manipulation did not alter EIT functioning. One interpretation of this finding is consistent with our hypothesis that the EIT is less amenable to contextual manipulations than other behavioral distress tolerance tasks. Of course, an alternative explanation is that the depletion manipulation was ineffective. Depletion paradigms may not be as robust as once thought (Inzlicht & Berkman, 2015; Lurquin et al., 2016), even though the "crossing out es" depletion method is well-cited in the literature (Hagger et al., 2010). In addition, we included an assessment of affect between the depletion manipulation and the EIT, which may have served a feelings-as-information function (Schwarz & Clore, 2003) whereby people had an affective explanation for their current state, thus weakening the depletion effect. Further research will be necessary to test whether the EIT is influenced by other contextual manipulations.

We also found some evidence that EIT persistence declines over time by assessing Block 1 compared to Block 5 performance. The decline was small (1-2 seconds, depending on the EIT index) but significant, suggesting that by the end of the EIT, people indicate distress earlier into the slide and are less willing to persist at viewing distressing images. This in and of itself could reflect a depletion effect, as continuing to enage with disturbing images may reflect diminishing self-control. Experts indicate that depletion and mental fatigue may be the same construct (Inzlicht & Berkman, 2015). It makes intuitive sense that as someone continues to engage with distressing images, they might grow mentally tired of "dealing with" them and thus rush to finish the task sooner. Our perspective is that this small shift in performance over the course of the EIT does not ruin the assessment of distress tolerance. In fact, it makes sense that trying to tolerate distress might become taxing over time, and people would become less adept at it the more they attempt to do it.

In addition, we found significant correlations between the EIT and both self-report measures of distress tolerance, although only the discomfort intolerance scale demonstrated a relationship with one of the persistence measures. Even so, greater discomfort intolerance and lower self-reported

distress tolerance were associated with shorter distress threshold and finding more of the slides distressing.

Finally, consistent with Study 2, we also found further evidence of EIT associations with binge eating symptoms, such that greater symptoms of binge eating were associated with lower overall image persistence, shorter distress threshold and finding more of the slides distressing. The current study, in contrast to Study 2, also found overlap between EIT performance and symptoms of depression. It may be that the correlations in this study were stronger because Study 2 also included the mirror tracing task and thus there was more noise in the study, or that the crossing out es task had some sort of influence on the results, even though not significantly different by depletion condition. Ultimately, these results provide additional verification that the EIT seems to capture something that is relevant to the study of psychopathology.

(Additional) References

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74*, 1252–1265. doi: 10.1037/0022-3514.74.5.1252

Gormally, J., Black S., Daston S., & Rardin D., C.D. (1982). The assessment of binge eating severity among obese persons. *Addictive Behaviors, 7*(1), 47-55.

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychological Bulletin, 136*(4), 495-525. doi: 10.1037/a0019486

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21: Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227-239. doi: 10.1348/014466505X29657

Inzlicht, M., & Berkman, E. (2015). Six questions for the resource model of control (and some answers). *Social and Personality Psychology Compass, 9*(10), 511-524. doi: 10.1111/spc3.12200

Lurquin, J. H., Michaelson, L. E., Barker, J. E., Gustavson, D. E., von Bastian, C. C., Carruth, N. P., & Miyake, A. (2016). No evidence of the ego-depletion effect across task characteristics and individual differences: A pre-registered study. *PloS one*, *11*(2), e0147770. doi: 10.1371/journal.pone.0147770

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle?. *Psychological Bulletin, 126*(2), 247 – 259. doi: 10.1037/0033-2909.126.2.247

Schmidt, N.B., Richey J.A., & Fitzpatrick, K.K. (2006). Discomfort intolerance: Development of a construct and measure relevant to panic disorder. *Journal of Anxiety Disorders*, *20(*3), 263-280. doi: 10.1016/j.janxdis.2005/02.2002

Schwarz, N., & Clore, G. (2003). Mood as information: 20 years later. *Psychological Inquiry, 14(*3), 296–303. doi:10.1207/S15327965PLI1403&4_20

Simons, J. S., & Gaher, R. M. (2005). The Distress Tolerance Scale: Development and validation of a self-report measure. *Motivation and Emotion, 29,* 83-102. doi: 10.1007/s11031-005-7955-3

Tice, D. M., Baumeister, R. F., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology, 43,* 379–384. doi: 10.1016/j.jesp.2006.05.007

Table 6. *Correlations between EIT and Study 4 variables*

	EIT Avg. Image Persistence	EIT Avg Distress Threshold	EIT Count Distress	EIT Avg Distress Persistence
NA Change EIT	-.20+	-.27*	.31**	-.11
Trait Distress Tolerance (DTS)	.14	.20*	-.23*	.03
Discomfort Intolerance (DIS)	-.30*	-.36**	.37**	-.14
Binge Eating (BES)	-.25*	-.30**	.29**	-.01
Depression (DASS)	-.24*	-.21+	.25*	-.20+
Anxiety (DASS)	.09	.05	-.02	-.01

+$p < .08$, *$p < .05$, **$p < .01$

EIT = Emotional Image Tolerance; DTS = Distress Tolerance Scale; DIS = Discomfort Intolerance Scale; BES = Binge Eating Scale; DASS = Depression, Anxiety & Stress Scales