

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

**Title:** Amazon Mechanical Turk as a platform for borderline personality disorder research

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

### **Abstract**

Researchers investigating the psychological processes underlying specific mental health problems often have difficulties achieving large enough samples for adequately powered studies. This can be particularly problematic when studying psychopathology with low base rates in typical samples (i.e., undergraduate and community). A relatively new approach to recruitment and testing employs online crowdsourcing to rapidly measure the characteristics and behavior of large numbers of people. We tested the feasibility of researching borderline personality disorder (BPD) in this manner using one large crowdsourcing site, Amazon Mechanical Turk (MTurk). Specifically, we examined prevalence rates of psychopathology in a large MTurk sample, as well as the demographic, psychosocial, and psychiatric characteristics of individuals who met criteria for BPD. These characteristics were compared across three groups: those who met criteria for BPD currently, those who met criteria for remitted BPD, and those who had never met criteria for BPD. The results suggest that MTurk may be ideally suited for studying individuals with a wide range of pathology, from healthy to intensely symptomatic to remitted.

**Keywords:** MTurk, crowdsourcing, borderline personality disorder, remission, recovery

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

## **Introduction**

The present study examines the feasibility of using Amazon Mechanical Turk (MTurk), an on-line data collection platform, to study borderline personality disorder (BPD). Specifically, we examine, across two studies, the correlates of BPD. We seek to recreate in data collected from MTurk participants the nomological network built around BPD in the literature, as well as prevalence rates established by current epidemiological research. We examine demographic, psychosocial, medical, and basic personality characteristics among three subgroups within our sample: those who meet criteria for current BPD, those who meet criteria for remitted BPD, and those who do not meet criteria for either. We evaluate the appropriateness of MTurk as a venue for studying BPD in the context of these findings.

## **MTurk**

MTurk is an online labor market where “workers” choose and complete small tasks, referred to as human intelligence tasks (HITs). Each task is listed with a description of the activity, time to complete, and proposed payment for completion, and workers are paid through the website after work is submitted and checked. HITs can be completed on a personal computer from anywhere.

There are several crowdsourcing markets, of which MTurk is currently the largest and most studied non-probability sample available to researchers (Chandler & Shapiro, 2016). MTurk is increasingly used in psychological research because it facilitates rapid collection of large datasets at relatively low cost (Chandler & Shapiro, 2016). Also, workers can be selected based on specific qualifications (e.g., native English-speakers, workers in the U.S., workers with high ratings of completion of other HITs), or pre-screened using a brief initial HIT to sort

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

subjects by other characteristics of interest prior to enrollment in a larger study.

MTurk is also attractive to researchers because it provides access to individuals who would not normally be represented in a clinic-based or convenience sample (Gosling & Mason, 2015). This group may include people with significant symptoms who have never presented for clinical care, and also those whose geographic locations, schedules, or preferences would dissuade them from presenting as research subjects. Demographics of MTurk workers are diverse, though not perfectly representative: workers are younger and better educated than the general population, and are primarily European or Asian American (Paolacci & Chandler, 2014). They have lower incomes and are more likely to be un- or under-employed (Corrigan, Bink, Fokuo, & Schmidt, 2015). Nonetheless, MTurk samples better match community samples than do typical research populations (e.g., college students and community volunteers; (Berinsky, Huber, & Lenz, 2012; Chandler & Shapiro, 2016).

Mental health problems are also well represented on MTurk as workers have anxiety and depression at levels comparable to the general population (Shapiro, Chandler, & Mueller, 2013). Additional research has examined personality disorders among MTurk workers, although this work has focused largely on narcissistic personality disorder. In fact, while the vast majority of research has examined NPD, only one of 98 studies examining personality disorders in MTurk published in two major psychology journals in the past three years focused on borderline personality disorder (Miller, Crowe, Weiss, Maples-Keller, & Lynam, 2017).

There are some limitations of the MTurk sample worth noting (Miller, Crowe, Weiss, Maples-Keller, & Lynam, 2017). Participants may be non-naïve, having potentially completed numerous psychological tasks on MTurk. They may be familiar with either the tasks or the methods used for checking attention and response validity. Also, the MTurk participant pool may

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

be smaller than the apparent millions of workers suggested by the platform: some data suggest that a small percentage of workers complete a substantial proportion of the HITs (Chandler, Mueller, & Paolacci, 2014). This requires researchers who hope to replicate or validate findings in repeated MTurk studies to attend closely to worker IDs in order to avoid recruiting duplicate individuals in subsequent samples.

### **BPD characteristics**

BPD is a disorder that is characterized by serious pervasive dysfunction and instability in personal relationships, affect, perceptual experience, and self-image. Impulsivity and intentional self-harm are common. BPD symptoms are associated with negative relationship experiences such as chronic relationship stress, partner dissatisfaction, relationship conflict, abuse, and unwanted pregnancy. In addition, adults with BPD report severe difficulty with multiple aspects of parenting, including the transition to parenthood, infant care, and early attachment (Zanarini, Frankenburg, Reich, Wedig, et al., 2015). Finally, given the substantial distress and impairment associated with this disorder, it is not surprising that it is associated with substantial health care utilization costs (Soeteman, Hakkaart-van Roijen, Verheul, & Busschbach, 2008).

### **BPD's Nomological Network**

#### **Prevalence and Comorbidities**

Data collected from a community sample of over 34,000 individuals in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) found that 5.9% of adults in the United States meet diagnostic criteria for BPD (Hasin & Grant, 2015). The NESARC found more frequent BPD among people who were female, low income, younger than 30,

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

separated or divorced, Native American or African American. Others have examined comorbidities in BPD: anxiety disorders, substance use disorders, and other personality disorders were most frequent (Swartz, Blazer, George, & Winfield, 1990; Tomko, Trull, Wood, & Sher, 2014). Lifetime prevalence of comorbid mood and trauma related disorders is 85%, prevalence of comorbid substance use disorders is 78% (Tomko, Trull, Wood, & Sher, 2014), and prevalence of comorbid chronic pain is 30% (Heath, Paris, Laporte, & Gill, 2017). BPD also increases risk for other medical problems (reviewed and discussed in (Iacovino, Powers, & Oltmanns, 2014).

### **Psychiatric medication use**

Zanarini and colleagues (Zanarini, Frankenburg, Reich, Harned, & Fitzmaurice, 2015) investigated psychiatric medication use in BPD compared to other personality disorders. They found that people with BPD took more antidepressants (1.3 times more), antipsychotics (2.6 times more) and mood stabilizers (3 times more) than did comparison subjects. Prescription opiate use was also more common in BPD than in other personality disorders (Frankenburg, Fitzmaurice, & Zanarini, 2014).

### **Self-harm**

Among inpatients with BPD, 90% endorse history of self-injury and more than 70% endorse history of suicide attempt(s) (Goodman et al., 2017). Other studies in adult samples have also found high rates of suicidal ideation and attempts among people with BPD (Wedig, Frankenburg, Reich, Fitzmaurice, & Zanarini, 2013; Wedig et al., 2012) (Paris & Zweig-Frank, 2001) (Boisseau et al., 2013) (Pompili, Girardi, Ruberto, & Tatarelli, 2005).

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

Between 5-10% complete suicide (Black, Blum, Pfohl, & Hale, 2004; Paris, 2006) (Pompili et al., 2005).

### **Basic personality traits**

Research has examined personality disorders as clusters of extreme variants of basic personality traits in the five factor model (FFM) (e.g., (Trull & Widiger, 2013) -- a classification system of normative personality traits that fall into five broad domains: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. In a recent meta-analysis, Samuel & Widiger (2008) found that BPD bears strong positive relations to neuroticism ( $r = .54$ ) and negative relations to agreeableness ( $r = -.24$ ) and conscientiousness ( $r = -.29$ ). These correlates are consistent with the FFM facet profile for BPD generated by experts (Lynam & Widiger, 2001). Similarly, (Hopwood & Zanarini, 2010) found that FFM neuroticism was significantly correlated to the affective ( $r = .50$ ), cognitive ( $r = .39$ ), and interpersonal domains ( $r = .45$ ) of BPD. In addition, FFM agreeableness was inversely correlated to impulsive symptoms ( $r = -.23$ ), and agreeableness and extraversion were inversely correlated to BPD symptoms overall.

### **Remission and recovery in BPD**

Research suggests that BPD symptoms remit over time for most people (75-99%) within a few years, and 40-60% achieve functional recovery (Zanarini, Frankenburg, Reich, & Fitzmaurice, 2012). Although these numbers are promising, people in remission from BPD are much less likely to exhibit high levels of functioning socially or in the community. For example, while people recovered from BPD are significantly more likely than non-recovered individuals

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

with BPD to have married or lived with a romantic partner, these same individuals are more likely to have given up or lost custody of a child (Zanarini, Frankenburg, Reich, Wedig, et al., 2015). These findings highlight the complexity of what it means to be in remission versus recovered and suggests that more research is needed to examine quality of life among these diagnostic groups.

### **Current Study**

The current study seeks to replicate, in an MTurk sample, the nomological network that surrounds BPD more broadly. We recruited more than 3,000 people via MTurk to answer questions about demographics, emotions, and behavior. Self-report surveys were administered including instruments to measure BPD symptoms now and in the past, as well as basic personality traits. A subset of this large group (~ 700 people) went on to a second experiment, which included measures of BPD, emotion regulation, impulsivity, depression and anxiety. We report on demographics and symptoms in people with current BPD, remitted BPD, and no history of BPD.

## **Methods**

### **Ethics**

This protocol and consent materials were approved by the [relevant institutional – specifics removed for masked file] Review Board. All participants were presented with the consent form online at the start of both parts one and two of the study. The consent information described anticipated risk (minimal risk due to potential discomfort when answering personal



**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

questions), confidentiality, a reminder that agreement to participate is voluntary and could be revoked, and contact information for the study team.

## **Participants**

Recruitment on MTurk was restricted to workers in the United States. Workers were invited to complete a 15-25 minute HIT about their “life, emotional state, and behavior” for \$0.40 compensation (HIT1). The study description also stated that some HIT1 participants would be invited to participate in a subsequent, more highly paid, HIT (HIT2 – see below). The MTurk site was linked to a survey platform on Qualtrics, where participants completed the surveys. Participants provided their MTurk ID and a survey completion code from Qualtrics to link their MTurk profile to their Qualtrics survey responses.

The initial group of HIT1 completers totaled 3,633 participants. We excluded participants for completing the task too quickly (< 9 minutes), for missing any of the three embedded attention check items, or for extreme frequency in responding (i.e. those who chose the same response choice at statistically outlying rates). After these exclusions, 3,132 participants remained. We also excluded people who reported color blindness, learning disability, traumatic brain injury, or schizophrenia. The ability to correctly perceive colors was important to a social cognition task included in HIT2, but not reported here. A final sample of 3,021 participants was used for analysis of HIT1 scales.

We invited a subset of HIT1 participants to participate in a second set of surveys and tasks, “HIT2”. We sorted the HIT1 completers into ten groups by gender (M/F) and BPD symptom level (based on SCID II questionnaire, endorsed currently having 0-2, 3-4, 5-9, 10+ symptoms, or having <5 current symptoms, but  $\geq$  5 past symptoms). For this process, we defined

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

current as within the past 2 years. HIT1 and HIT2 ran simultaneously: as people completed HIT2, we aimed to fill each of the 10 categories with 30 participants. When each category was full, we stopped inviting HIT1 participants for that category. This approach allowed us to oversample for BPD symptoms and for male participants in HIT2, but it does result in a HIT2 sample that is not representative of the general MTurk population. This process resulted in an initial sample of 1,134 participants, all of whom had completed HIT1. These categories were used only for recruitment. Please see below in the Methods section regarding the SCID-II-PQ-BPD scale for the definitions of the three groups we defined for analysis.

HIT2 participants experienced a consent process as described for HIT1, however it described a longer HIT of ~ 40 minutes, including social games, and compensation of \$2.00 plus a possible \$2.00 bonus for excellent game performance. We excluded people who were ultimately excluded from HIT1 analyses (due to extreme responses, etc.;  $n = 243$ ). We then excluded people who did not pass HIT2 attention checks ( $n = 52$ ), or missed some survey questions ( $n = 128$ ). A final sample of 711 participants was used for analysis of HIT2 surveys.

## **HIT1 Measures**

**Demographics and medical history questionnaire.** Participants responded to a series of questions asking about demographic characteristics, past medical and psychiatric history, and current medications. Medications were manually read and divided into categories by psychiatrists (SKF and ER).

**International Personality Item Pool Representation of the NEO PI-R Short Form (IPIP-NEO SF).** The short form of the IPIP-NEO (Maples-Keller, et al., 2017) is comprised of

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

60 items that measure the thirty facets of the five domains (neuroticism, openness, conscientiousness, agreeableness, and extraversion) of the Five Factor Model (FFM). Twelve items comprise each domain, with two items representing each of the thirty FFM facets. In the present study, alpha coefficients for the domain scales ranged from .57 to .77.

**Structured Clinical Interview for DSM-IV Axis II Personality Disorders Self-report Questionnaire (SCID-II-PQ-BPD).** We employed the 15 BPD questions from the SCID-II personality disorders self-report questionnaire, which assess the nine diagnostic criteria for BPD as outlined in DSM-IV (First, Gibbon, Spitzer, Williams, & Benjamin, 1997; First, Spitzer, Gibbon, Williams, Benjamin, 1994). For each item that asked participants if they had experienced a given symptom, they were given four response choices: “Yes, that happened this month”; “Not this month, but in the last 2 years”; “Not in the last 2 years, but in the past”; and “No.” The alpha coefficient for the SCID-II-PQ-BPD in the current study was .88.

**Participant Group Assignments for Analysis**

We used the SCID-II-PQ-BPD responses to define three groups. First, the fifteen questions were used to code binary responses to each of the nine DSM criteria. If participants reported at least five of the nine DSM criteria in the last two years (we combined responses “yes, that happened this month” and “not this month, but in the last 2 years”), they were assigned to the group “Current BPD” (n = 1,020). If participants reported at least five of the nine criteria in their lifetime, but not within the last two years, they were assigned to the “Remitted BPD” group (n = 865). The remaining participants were assigned to the “Never BPD” group (n = 1,136).

Participants in HIT1 also completed several tasks that are not discussed in this paper.

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

These were measures of psychopathy, narcissism, autism, a brief writing sample in response to the prompt “Please tell us about yourself,” and a five-round Trust Game.

## **HIT2 Measures**

**Beck Anxiety Inventory (BAI).** The BAI is a 21-item self-report measure assessing various symptoms of anxiety (Beck, Epstein, Brown, & Steer, 1988; Fydrich, Dowdall, & Chambless, 1992). Participants rate the symptoms (e.g., “Numbness or tingling”) on a scale from 0 (Not at all) to 3 (Severely - it bothered me a lot) indicating how much he/she has been bothered by that symptom during the past month, with a total possible score of 63 points. The BAI demonstrated good reliability in the present study ( $\alpha = .94$ ).

**Borderline Symptom List 23 (BSL-23).** The BSL-23 consists of 23 items assessing BPD symptoms experienced in the past week (Bohus et al., 2009). Participants indicate how the extent to which they experienced each item on a scale from 0 (Not at all) to 4 (Very strong). Sample items include “I was lonely” and “I suffered from shame.” In the current study, the BSL-23 had a coefficient alpha of .95.

**World Health Organization Quality of Life - BREF (WHOQOL-BREF).** The WHOQOL-BREF (The WHOQOL Group, 1998) is a 26-item scale designed to assess perceptions of physical health, psychological health, social relationships, and environment. Item responses are given on a five-point Likert scale. Reliability for this scale was .93.

**Beck Depression Inventory Second Edition (BDI-II).** The BDI-II (Beck, Steer, &

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

Brown, 1996) consists of 21 items and is used to assess the existence and severity of depression symptoms in the past two weeks as outlined in the DSM-IV. There is a four-point scale for each item ranging from 0 to 3, with higher scores indicating elevated symptom levels. The BDI demonstrated strong reliability in the present study ( $\alpha = .95$ ).

**Barratt Impulsiveness Scale (BIS-11).** The BIS-11 (Patton, Stanford, & Barratt, 1995) was designed to assess multiple facets of impulsivity. The BIS-11 comprises 30 items, all on a 4-point scale ranging from “Rarely/Never” to “Almost Always/Always.” Reliability for this was measure was .86 overall and .65-.76 for the subscales in the current study.

**Difficulties in Emotion Regulation Scale (DERS).** The DERS (Gratz & Roemer, 2004) is a 36-item measure that assesses a person’s beliefs and feelings when he/she becomes upset, as well as a person’s perceptions of how they experience emotion. Items are measured on a five-point Likert scale. Coefficient alpha was 0.87.

**Positive and Negative Affect Scale (PANAS).** The PANAS (Watson, Clark, & Tellegen, 1988) comprises two mood scales, one assessing positive affect and the other assessing negative affect. There are 20 items in the measure, each on a five-point Likert scale. Both the positive and negative affect scales demonstrated good reliability ( $\alpha = .92$  and  $\alpha = .89$ , respectively).

After responding to these self report measures, HIT2 participants also completed two tasks not discussed here: one 10 round Cyberball game (Williams & Jarvis, 2006; Williams, Yaeager, Cheung, & Choi, 2012) followed by an associative learning task focused on social cognition.

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

## **Results**

### **Characteristics of the overall sample**

The demographics of our sample (HIT1) were consistent with past research on the MTurk population (Berinsky, Huber, & Lenz, 2012; Shapiro, 2013; Buhrmester, Kwang, & Gosling, 2011; Paolacci & Chandler, 2014). Participants were overwhelmingly female (66.1%), white (74%), middle-aged ( $M = 37.12$  years, range 18 – 81,  $SD = 12.26$ ), heterosexual (84.9%), and well educated (86.2% had completed at least some college). Only a small fraction (20.4%) were currently in school, but most (70.3%) were employed at least 10-20 hours/week, and 43.2% were employed full-time.

### **Demographics of people with current, remitted, and never BPD**

We found that 1020 HIT1 participants (33.8%) met DSM criteria for BPD in the last 2 years – we termed this group “**Current BPD**”. Another 28.6% of HIT1 participants endorsed having had 5+ DSM criteria for BPD sometime during their life, but having  $\leq 5$  DSM BPD criteria in the past two years. This group was termed “**Remitted BPD.**” The remaining participants, who had fewer than five lifetime BPD symptoms, belong to the “**Never BPD**” group. Of note, only 1% of the overall sample, and 2.2% of the Current BPD group had been previously diagnosed with BPD (**Table 1**).

We observed some demographic differences among the three groups (**Table 1**). People with BPD were younger (Never BPD  $M = 41.03$ ,  $SD = 13.48$ ; Remitted BPD  $M = 37.87$ ,  $SD = 11.67$ ; Current BPD  $M = 32.17$ ,  $SD = 9.27$ ,  $F(2, 2989) = 155.70$ ,  $p < 0.001$ , post hoc for all 3

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

comparisons  $p < 0.001$ ). There were no significant group differences by gender (Never BPD = 66.6% female, Remitted BPD = 63.4% female, Current BPD = 67.9% female,  $X^2(2) = 4.49, p = 0.11$ ).

**Co-morbidities of people with current, remitted, and never BPD**

We examined common co-morbidities in people with BPD (**Table 1**). As expected, people with BPD were more likely to endorse anxiety ( $X^2(2) = 284.72, p < 0.001$ , Cramer's  $V = 0.307$ ), depression ( $X^2(2) = 295.42, p < 0.001$ , Cramer's  $V = 0.313$ ), and bipolar disorder ( $X^2(2) = 76.43, p < 0.001$ , Cramer's  $V = 0.159$ ). For anxiety, depression, and bipolar disorder, the Remitted BPD represented an intermediate phenotype (post-hoc tests for all 3 comparisons significant at  $p < 0.05$ ).

Participants with BPD were also more likely to endorse substance use disorders ( $X^2(2) = 33.06, p < 0.001$ , Cramer's  $V = 0.105$ ). Of note, in the overall sample, 3-4% of participants reported a diagnosis of substance use disorder (SUD), which is comparable to previous work on MTurk, where 4.3% endorsed having sought treatment for SUD (though 37.1% endorsed SUD-consistent symptoms) (Shapiro, Chandler, & Mueller, 2013). We did not use self-reports to screen for SUD symptoms here, we only asked about diagnosis. Here, the remitted group reported significantly more frequent SUD diagnosis than Never BPD participants ( $p < 0.05$ ), and this was not significantly different than Current BPD group frequency.

Chronic pain was also more frequent in Current BPD than in Never BPD (23.4% versus 17.3%,  $X^2(2), p < .01$ , Cramer's  $V = 0.065$ ), however frequency in Remitted BPD was not significantly different than either of the other two groups.

People with BPD were more likely to be taking psychotropic medications, including

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.**

**Please do not copy or cite without author's permission.**

antidepressants ( $X^2(2) = 76.53, p < 0.001$ , Cramer's  $V = 0.159$ ), mood stabilizers ( $X^2(2) = 34.69, p < 0.001$ , Cramer's  $V = 0.107$ ), antipsychotics ( $X^2(2) = 14.02, p < 0.001$ , Cramer's  $V = 0.068$ ), benzodiazepines ( $X^2(2) = 17.1, p < 0.001$ , Cramer's  $V = 0.075$ ), and other miscellaneous psychotropics ( $X^2(2) = 12.29, p < .001$ , Cramer's  $V = 0.064$ ). For anti-depressants and mood-stabilizers, frequency is at an intermediate level in the Remitted BPD group (significant differences between each of the 3 groups,  $p < 0.05$ ). We did not observe significant differences in opiate or muscle relaxant use between groups.

### **Relationship status of people with current, remitted, and never BPD**

We found significant differences in both current relationship status ( $X^2(2) = 137.33, p < 0.001$ , Cramer's  $V = 0.151$ ) and parental status ( $X^2(2) = 50.31, p < 0.001$ , Cramer's  $V = 0.129$ ) in our sample (**Table 2**). Participants in the Current BPD group were less likely to be married (51.8% of Never BPD, 32.5% of Current BPD), and more likely to be in a relationship but not married (16.2% of Never BPD, 28.7% of Current BPD) or single (21.3% of Never BPD, 30.6% of Current BPD). Participants in the Remitted BPD group demonstrated an intermediate phenotype for frequency of participants married (46.4%) and in a relationship, not married (23.5%). Likelihood of being a parent decreases from Never BPD to Remitted BPD to Current BPD.

### **BPD symptoms endorsed in HIT1 and HIT2 surveys**

We also examined the number and specifics of the BPD symptoms endorsed on the HIT1 assessment (SCID-II self-report questionnaire) (**Table 3**). People in the Current BPD group endorsed an average of 6.57 current symptoms and 1.85 remitted symptoms. People in the



**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

Remitted BPD endorsed an average of 2.78 current symptoms and 3.69 remitted symptoms.

People in the Never BPD group endorsed 1.19 current symptoms and 1.19 remitted symptoms.

We tested our participants' endorsement of BPD symptoms for reliability by measuring the correlation between score on the BPD symptom report in HIT1 (SCID-II self-report questionnaire) and BPD symptom report in HIT2 (BSL-23), which occurred days to weeks later (mean = 16.8 days). The number of self-reported symptoms of BPD on the SCID-II and the total score on the BSL-23 were positively correlated at  $r = .58, p < .001$ .

### **Psychiatric co-morbidities by group**

In HIT2, we examined symptoms and personality traits of participants in each group. (**Table 3: HIT2 Scales**). We found the expected significant between-group differences, with Current BPD > Remitted BPD > Never BPD on all scales (post-hoc tests all  $p < 0.05$ ). Specifically, people with BPD endorsed more BPD symptoms ( $F(2, 3018) = 135.21, p < 0.001, \eta^2 = 0.295$ ), more difficulties with emotion regulation ( $F(2, 550) = 151.46, p < 0.001, \eta^2 = 0.301$ ), more anxiety ( $F(2, 493) = 112.5, p < 0.001, \eta^2 = 0.272$ ), more depression ( $F(2, 537) = 133.05, p < 0.001, \eta^2 = 0.297$ ), more impulsivity ( $F(2, 476) = 81.20, p < 0.001, \eta^2 = 0.207$ ), and more negative affect ( $F(2, 543) = 60.92, p < 0.001, \eta^2 = 0.134$ ) than did people without BPD. Also, people with BPD endorsed less positive affect ( $F(2, 545) = 39.25, p < 0.001, \eta^2 = 0.074$ ) and lower quality of life (subscales with F-tests ranging from  $F(2, 550) = 41.47 - 102.62, p < 0.001, \eta^2 = 0.131 - 0.272$ ) than did people without BPD.

We also observed the expected correlations between symptom scale scores (**Table S1**).

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

### **Personality traits of participants with current, remitted, and never BPD**

Participants with BPD also reported personality traits consistent with past reports (**Table 3**). These data, collected in HIT1, included the full sample of 3,021 individuals, and we found significant differences (post-hoc tests  $p < 0.001$ ) between each of the three groups for each of the big five traits (Current BPD vs. Remitted BPD vs. Never BPD). People with BPD endorsed higher trait neuroticism,  $F(2, 3018) = 717.86, p < 0.001, \eta^2 = 0.322$ , and openness,  $F(2, 3018) = 45.16, p < 0.001, \eta^2 = 0.029$ , and lower extraversion,  $F(2, 3018) = 63.43, p < 0.001, \eta^2 = 0.040$ , agreeableness,  $F(2, 3018) = 63.72, p < 0.001, \eta^2 = 0.041$ , and conscientiousness,  $F(2, 3018) = 296.36, p < 0.001, \eta^2 = 0.164$ .

### **History of self-harm in participants with current, remitted, and never BPD**

An examination of the chronology of self-harm in each group revealed significant and large between-group differences (**Figure 1**, tested 3 groups x 3 time frames (lifetime, past, in the last month):  $X^2(4) = 602.34, p < 0.001$ , Cramer's  $V = 0.316$ ). Lifetime history of self-harm was more frequent in people with BPD (Never BPD 4.7%, Remitted BPD 27.3%, Current BPD 50%, and post-hocs revealed significant differences between each of the 3 groups ( $p < 0.05$ ; **Figure 1A**). Past history of self-harm was also more frequent in people with BPD (Never BPD 4.6%, Remitted BPD 27.1%, Current BPD 44.6%, and post-hoc tests revealed significant differences between each of the 3 groups ( $p < 0.05$ ; **Figure 1B**). Finally, past month history of self-harm was also more frequent in people with BPD (Never BPD 0.10%, Remitted BPD 0.20%, Current BPD 5.50%, however, here post-hocs distinguished only the Current BPD group from the other two groups ( $p < 0.05$ ; **Figure 1C**).

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

## **Discussion**

In this study, we aimed to test MTurk as a potential venue for BPD research. We compared prevalence and associated factors in a large MTurk sample to previously established data from community and patient samples. We tested for respondents' careful completion of surveys and reliability of their responses. We measured demographics and frequency of past mental health diagnosis and of self-reported personality traits and BPD symptoms now and in the past. We also measured these features in the Remitted BPD group. Overall, we found that BPD features are quite frequent among MTurk workers, both in terms of direct interrogation of DSM criteria, and expected features of social life, symptoms, and personality.

### **Overall sample**

Our sample was similar to other researchers' MTurk samples in terms of demographics and self-reported psychopathology (Berinsky et al., 2012; Buhrmester et al., 2011; Paolacci & Chandler, 2014; Shapiro, Chandler, & Mueller, 2013). With regard to BPD symptoms in particular, we found a very high frequency of self-reported symptoms, with one-third of our participants meeting DSM criteria for BPD. This is far higher than other large community samples that have found prevalence rates to be closer to 5-6% (Grant et al., 2008).

We tested for consistency and reliability of responding, and found high correlations between scores on our two different BPD self-reports, and between BPD scores and mood and emotional regulation symptom scales. We also found that BPD scores correlated with lower self-reported quality of life, lower positive affect, higher negative affect, and higher neuroticism. Meeting criteria for BPD in our sample also predicted more psychiatric medication use and

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

higher rates of obesity and high blood pressure. These results are consistent with the expected profile for people with BPD.

There are several possible explanations for this surprisingly high frequency of people with BPD in our sample, in addition to the higher base rates typically yielded by self-report measures (Hopwood et al., 2008). MTurk may be attractive to people with BPD, as it allows for social interactions with boundaries set by task directions, durations, and contracts, but also allows the MTurk worker a great deal of flexibility in aspects of engagement such as where to work, when to work, how long to work, what to work on, and who to work with. This combination of clear social structure and personal flexibility may facilitate participation by this group of people with high anxiety, fluctuating symptoms, and social difficulties.

Another consideration is whether our symptom measures are specific tests for BPD or more general markers of psychopathology. We did observe high correlations between BPD, depression, and anxiety self-report measures, consistent with known patterns of co-morbidity in BPD. We also note that in our sample, among people who met criteria for current BPD, 50% had at some point in their lives tried to or threatened to hurt or kill themselves, consistent with chronic suicidality and marked increases in the frequency of suicidal action in BPD.

Additionally, although self-report methods of obtaining diagnoses cannot replace a diagnosis given by a clinician following a structured diagnostic interview, the SCID-II-PQ has been found to be sensitive and specific, with moderate-to-excellent accuracy when compared to clinical interviews (Fowler et al., 2018).

**Discrepancy between high frequency of meeting BPD criteria and low frequency of past diagnosis**

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

Only a small number of those meeting criteria for BPD reported a previous diagnosis of BPD, though many carried diagnoses of mood and anxiety disorders. The discrepancy between symptom-reports and previous diagnosis may arise because of under-diagnosis or misdiagnosis in the community. This is a frequent problem, likely driven by lack of clinician experience in BPD diagnosis, or preference for giving a diagnosis with less stigma that is perceived to be more treatable (e.g. (Gunderson, 2009).

### **Remission**

Our approach to defining a remitted BPD group here depends on the participant's retrospective recall of their past symptoms. This is a major limitation of this approach. However, report of current symptoms and social status cohere with the claim that people in this group are at an intermediate level of difficulty. We hope that future work on BPD in this setting will probe remission and recovery in more detail to better understand this group. We think that both the Current and Remitted BPD groups here likely include individuals who, on detailed clinical interview, may be more in the BPD traits (< 5 DSM criteria) range than the full syndrome range. However, some people may also under-report symptoms to a novel interviewer but feel comfortable in the potentially more anonymous feeling setting of an online survey. Identifying how life experience changes as BPD symptoms are ameliorated, and what factors may predict and promote recovery, are critical questions. The MTurk worker population does appear to include a large number of people with high and mid-range levels of BPD symptoms.

### **Potential implications of findings**

We have defined the frequency and characteristics of MTurk workers who meet criteria

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.**

**Please do not copy or cite without author's permission.**

for current or remitted BPD. Our findings suggest that MTurk will be a good locus for future work to test survey response and potentially also cognitive tasks in people with BPD. Large sample sizes in this setting hold promise for doing studies that provide more statistical power, better engage a dimensional approach, and that consider the experiences of people in remission and in recovery.

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

**References**

- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol*, 56(6), 893-897.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Berinsky, A., Huber, G., & Lenz, G. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3), 351-368. doi:10.1093/pan/mpr057
- Black, D. W., Blum, N., Pfohl, B., & Hale, N. (2004). Suicidal behavior in borderline personality disorder: prevalence, risk factors, prediction, and prevention. *J Pers Disord*, 18(3), 226-239. doi: 10.1521/pedi.18.3.226.35445
- Bohus, M., Kleindienst, N., Limberger, M. F., Stieglitz, R. D., Domsalla, M., Chapman, A. L., . . . Wolf, M. (2009). The short version of the Borderline Symptom List (BSL-23): development and initial data on psychometric properties. *Psychopathology*, 42(1), 32-39. doi: 10.1159/000173701
- Boisseau, C. L., Yen, S., Markowitz, J. C., Grilo, C. M., Sanislow, C. A., Shea, M. T., . . . McGlashan, T. H. (2013). Individuals with single versus multiple suicide attempts over 10years of prospective follow-up. *Compr Psychiatry*, 54(3), 238-242. doi: 10.1016/j.comppsy.2012.07.062
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspect Psychol Sci*, 6(1), 3-5. doi: 10.1177/1745691610393980
- Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaivete among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers. *Behav Res Methods*, 46(1), 112-130. doi: 10.3758/s13428-013-0365-7
- Chandler, J., & Shapiro, D. (2016). Conducting Clinical Research Using Crowdsourced Convenience Samples. *Annu Rev Clin Psychol*, 12, 53-81. doi: 10.1146/annurev-clinpsy-021815-093623
- Corrigan, P. W., Bink, A. B., Fokuo, J. K., & Schmidt, A. (2015). The public stigma of mental illness means a difference between you and me. *Psychiatry Res*, 226(1), 186-191. doi: 10.1016/j.psychres.2014.12.047
- First, M. B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., Benjamin, L.S. (1997). User's Guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II). Washington, D.C.: American Psychiatric Press.
- First, M. B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., Benjamin, L. (1994). Structured Clinical Interview for DSM-IV axis II personality disorders (SCID-II) (Version 2.0). In N. Y. S. P. I. Biometric Research Department (Ed.). New York.
- Fowler, J.C., Madan, A., Allen, J.G., Patriquin, M., Sharp, C., Oldham, J.M., Frueh, B.C. (2018). Clinical Utility of the DSM-5 alternative model for borderline personality disorder: Differential diagnostic accuracy of the BFI, SCID-II PQ, and PID-5. *Compr Psychiatry*, 80:97-103. doi: 10.1016/j.comppsy.2017.09.003
- Frankenburg, F. R., Fitzmaurice, G. M., & Zanarini, M. C. (2014). The use of prescription opioid medication by patients with borderline personality disorder and axis II comparison subjects: a 10-year follow-up study. *J Clin Psychiatry*, 75(4), 357-361. doi: 10.4088/JCP.13m08557

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.**

**Please do not copy or cite without author's permission.**

- Fydrich, T., Dowdall, D., & Chambless, D. L. (1992). Reliability and Validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders*, 6(1), 55-61. doi: Doi 10.1016/0887-6185(92)90026-4
- Goodman, M., Tomas, I. A., Temes, C. M., Fitzmaurice, G. M., Aguirre, B. A., & Zanarini, M. C. (2017). Suicide attempts and self-injurious behaviours in adolescent and adult patients with borderline personality disorder. *Personal Ment Health*, 11(3), 157-163. doi: 10.1002/pmh.1375
- Gosling, S. D., & Mason, W. (2015). Internet research in psychology. *Annu Rev Psychol*, 66, 877-902. doi: 10.1146/annurev-psych-010814-015321
- Grant, B. F., Chou, S. P., Goldstein, R. B., Huang, B., Stinson, F. S., Saha, T. D., . . . Ruan, W. J. (2008). Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*, 69(4), 533-545.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, 26, 41-54.
- Gunderson, J. G. (2009). Borderline personality disorder: ontogeny of a diagnosis. *Am J Psychiatry*, 166(5), 530-539. doi: 10.1176/appi.ajp.2009.08121825
- Hasin, D. S., & Grant, B. F. (2015). The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Soc Psychiatry Psychiatr Epidemiol*, 50(11), 1609-1640. doi: 10.1007/s00127-015-1088-0
- Heath, L. M., Paris, J., Laporte, L., & Gill, K. J. (2017). High Prevalence of Physical Pain Among Treatment-Seeking Individuals With Borderline Personality Disorder. *J Pers Disord*, 1-7. doi: 10.1521/pedi\_2017\_31\_302
- Hopwood, C. J., & Zanarini, M. C. (2010). Borderline personality traits and disorder: predicting prospective patient functioning. *J Consult Clin Psychol*, 78(4), 585-589. doi: 10.1037/a0019003
- Iacovino, J. M., Powers, A. D., & Oltmanns, T. F. (2014). Impulsivity mediates the association between borderline personality pathology and body mass index. *Pers Individ Dif*, 56, 100-104. doi: 10.1016/j.paid.2013.08.028
- Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach. *Journal of Abnormal Psychology*, 110(3), 401-412.
- Maples-Keller, J. L., Williamson, R. L., Sleep, C. E., Carter, N. T., Campbell, W. K., & Miller, J. D. (2017). Using Item Response Theory to Develop a 60-item version Representation of the NEO- PI-R using the International Personality Item Pool: Development of the IPIP-NEO-60. *Journal of Personality Assessment*, 1-12.
- Miller, J. D., Crowe, M., Weiss, B., Maples-Keller, J. L., & Lynam, D. R. (2017). Using online, crowdsourcing platforms for data collection in personality disorder research: The example of Amazon's Mechanical Turk. *Personal Disord*, 8(1), 26-34. doi: 10.1037/per0000191
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23(3), 184-188.



**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

- Paris, J. (2006). Managing suicidality in patients with Borderline Personality Disorder. *Psychiatric Times*. Retrieved from <http://www.psychiatristimes.com/articles/managing-suicidality-patients-borderline-personality-disorder>
- Paris, J., & Zweig-Frank, H. (2001). A 27-year follow-up of patients with borderline personality disorder. *Compr Psychiatry*, 42(6), 482-487. doi: 10.1053/comp.2001.26271
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, 51(6), 768-774.
- Pompili, M., Girardi, P., Ruberto, A., & Tatarelli, R. (2005). Suicide in borderline personality disorder: a meta-analysis. *Nord J Psychiatry*, 59(5), 319-324. doi: 10.1080/08039480500320025
- Samuel, D.B. & Widiger, T.A. (2008). A Meta-Analytic Review of the Relationships Between the Five-Factor Model and DSM-IV-TR Personality Disorders: A Facet Level Analysis. *Clin Psychol Rev*, 28(8), 1326-1342.
- Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using Mechanical Turk to study clinical populations. *Clinical Psychological Sciences*, 1(2), 213-220.
- Soeteman, D. I., Hakkaart-van Roijen, L., Verheul, R., & Busschbach, J. J. (2008). The economic burden of personality disorders in mental health care. *J Clin Psychiatry*, 69(2), 259-265.
- Swartz, M., Blazer, D., George, L., & Winfield, I. (1990). Estimating the prevalence of borderline personality disorder in the community. *J Pers Disord*, 4(3), 257-272.
- The WHOQOL Group. (1998). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med*, 28(3), 551-558.
- Tomko, R. L., Trull, T. J., Wood, P. K., & Sher, K. J. (2014). Characteristics of borderline personality disorder in a community sample: comorbidity, treatment utilization, and general functioning. *J Pers Disord*, 28(5), 734-750. doi: 10.1521/pedi\_2012\_26\_093
- Trull, T. J., & Widiger, T. A. (2013). Dimensional models of personality: the five-factor model and the DSM-5. *Dialogues Clin Neurosci*, 15(2), 135-146.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol*, 54(6), 1063-1070.
- Wedig, Frankenburg, F., Reich, G., Fitzmaurice, G., & Zanarini, M. (2013). Predictors of suicide threats in patients with borderline personality disorder over 16 years of prospective follow-up. *Psychiatry Res*, 208(3), 252-256. doi: 10.1016/j.psychres.2013.05.009
- Wedig, M., Silverman, M., Frankenburg, F., Reich, D., Fitzmaurice, G., & Zanarini, M. (2012). Predictors of suicide attempts in patients with borderline personality disorder over 16 years of prospective follow-up. *Psychol Med*, 42(11), 2395-2404. doi: 10.1017/S0033291712000517
- Williams, K. D., & Jarvis, B. (2006). Cyberball: a program for use in research on interpersonal ostracism and acceptance. *Behav Res Methods*, 38(1), 174-180.
- Williams, K. D., Yeaeger, D.S., Cheung, C.K.T., Choi, W. (2012). Cyberball (version 4.0). Retrieved from <http://cyberball.wikispaces.com>
- Zanarini, M.C., Frankenburg, F., Reich, B., Harned, A., & Fitzmaurice, G. (2015). Rates of psychotropic medication use reported by borderline patients and axis II comparison subjects over 16 years of prospective follow-up. *J Clin Psychopharmacol*, 35(1), 63-67. doi: 10.1097/JCP.0000000000000232
- Zanarini, M. C., Frankenburg, F. R., Reich, D. B., & Fitzmaurice, G. (2012). Attainment and stability of sustained symptomatic remission and recovery among patients with borderline

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.**

**Please do not copy or cite without author's permission.**

personality disorder and axis II comparison subjects: a 16-year prospective follow-up study. *Am J Psychiatry*, 169(5), 476-483.

Zanarini, M. C., Frankenburg, F. R., Reich, D. B., Wedig, M. M., Conkey, L. C., & Fitzmaurice, G. M. (2015). The course of marriage/sustained cohabitation and parenthood among borderline patients followed prospectively for 16 years. *J Pers Disord*, 29(1), 62-70. doi: 10.1521/pedi\_2014\_28\_147

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

**Table 1.** Previous diagnoses and medications in HIT1 sample (N = 3,021).

	Never BPD n = 1136	Remitted BPD n = 865	Current BPD n = 1020	Pearson Chi-square	Sig.
<b>Previous Diagnoses</b>					
Borderline PD	0.00% <sup>a</sup>	0.80% <sup>b</sup>	2.20% <sup>b</sup>	26.59, df = 2	p < 0.001
Depression	14.40% <sup>a</sup>	32.50% <sup>b</sup>	48.70% <sup>c</sup>	295.42, df = 2	p < 0.001
Anxiety	17.20% <sup>a</sup>	34.70% <sup>b</sup>	51.60% <sup>c</sup>	284.72, df = 2	p < 0.001
Bipolar Disorder	0.40% <sup>a</sup>	3.20% <sup>b</sup>	7.50% <sup>c</sup>	76.43, df = 2	p < 0.001
Substance Use Disorder	1.10% <sup>a</sup>	3.60% <sup>b</sup>	5.60% <sup>b</sup>	33.06, df = 2	p < 0.001
Chronic pain	17.30% <sup>a</sup>	20.90% <sup>ab</sup>	23.40% <sup>b</sup>	12.85, df = 2	p = 0.002
<b>Current Medications</b>					
Antidepressant	8.50% <sup>a</sup>	17.50% <sup>b</sup>	22.00% <sup>c</sup>	76.53, df = 2	p < 0.001
Mood stabilizer	0.50% <sup>a</sup>	1.40% <sup>b</sup>	3.90% <sup>c</sup>	34.69, df = 2	p < 0.001
Antipsychotic	0.30% <sup>a</sup>	0.90% <sup>b</sup>	1.90% <sup>b</sup>	14.02, df = 2	p = 0.001
Benzodiazepine	1.80% <sup>a</sup>	3.70% <sup>b</sup>	5.10% <sup>b</sup>	17.1, df = 2	p < 0.001
Other psychotropics	3.30% <sup>a</sup>	3.90% <sup>a</sup>	6.30% <sup>b</sup>	12.29, df = 2	p = 0.002
Opiate	1.80% <sup>a</sup>	2.70% <sup>a</sup>	1.70% <sup>a</sup>	2.60, df = 2	p = 0.273
Muscle relaxant	0.50% <sup>a</sup>	1.20% <sup>a</sup>	0.60% <sup>a</sup>	3.09, df = 2	p = .213

*Notes.* For each variable, we present the % of the group (Never, Remitted, or Current BPD) who endorse that variable. Post hoc test for ANOVAs is Bonferroni, z-test was used to compare column proportions after Chi-square. Superscript letters that differ across comparison groups indicate statistically significant difference,  $p < .05$ . Any groups that have the same letters did not statistically differ from each other.

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

**Table 2.** Relationship status of full HIT1 sample (N = 3,021).

	<b>N: Never BPD (% of group)</b>	<b>R: Remitted BPD (% of group)</b>	<b>B: Current BPD (% of group)</b>	<b>Pearson chi square</b>	<b>Sig.</b>
<b>Current Relationship Status</b>				137.328, df = 12	p < 0.001
Married	51.80% <sup>a</sup>	46.40% <sup>b</sup>	32.50% <sup>c</sup>		
In a relationship, not married	16.20% <sup>a</sup>	23.50% <sup>b</sup>	28.70% <sup>c</sup>		
Divorced	6.80% <sup>a</sup>	4.70% <sup>b</sup>	4.10% <sup>b</sup>		
Separated	1.10% <sup>a</sup>	0.70% <sup>a</sup>	1.30% <sup>a</sup>		
Single	21.30% <sup>a</sup>	22.40% <sup>a</sup>	30.60% <sup>b</sup>		
Widowed	1.70% <sup>a</sup>	0.30% <sup>b</sup>	0.30% <sup>b</sup>		
Other	1.20% <sup>a</sup>	2.00% <sup>a</sup>	2.50% <sup>a</sup>		
<b>Parent</b>	58.50% <sup>a</sup>	52.80% <sup>b</sup>	43.30% <sup>c</sup>	50.313, df = 2	p < 0.001

*Note.* Post hoc test for ANOVAs is Bonferroni, z-test was used to compare column proportions after Chi-square. Superscript letters that differ across comparison groups indicate statistically significant difference,  $p < .05$ . Any groups that have the same letters did not statistically differ from each other.

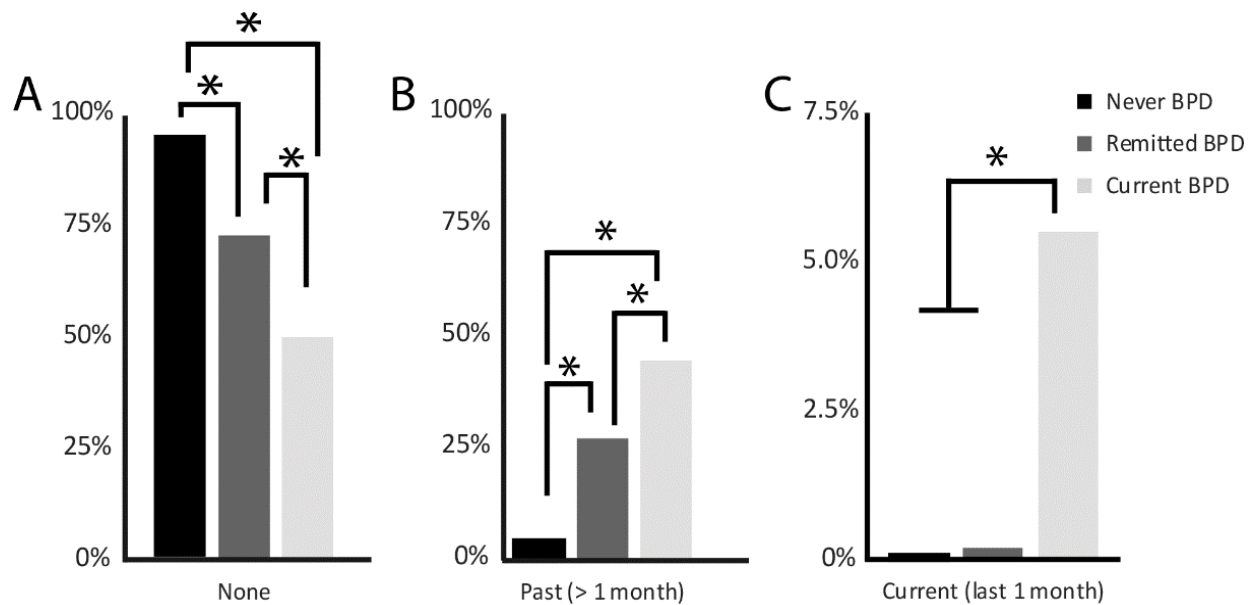
**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

**Table 3.** BPD symptoms across diagnostic subgroups. Data presented as mean value, standard deviation.

	Never BPD	Remitted BPD	Current BPD	One-way ANOVA	Sig.
<b>Current DSM BPD symptoms (past 2 years)</b>	1.19, 1.09 <sup>a</sup>	2.78, 1.19 <sup>b</sup>	6.57, 1.32 <sup>c</sup>	F = 5592.21, df = 2	p < 0.001
<b>Remitted DSM BPD symptoms</b>	1.19, 1.10 <sup>a</sup>	3.69, 1.62 <sup>b</sup>	1.85, 1.39 <sup>c</sup>	F = 859.51, df = 2	p < 0.001
<b>HIT1 to HIT2 reliability</b>	<b>Time elapsed</b>			<b>Pearson correlation</b>	<b>Sig.</b>
<b>DSM-BPD sx to BSL23 score</b>	16.80 days, 24.42			0.582	p < 0.001
	Never BPD n = 255	Remitted BPD n = 180	Current BPD n = 276	One-way ANOVA	Sig.
<b>HIT2 scales</b>					
<b>BPD symptoms</b>	7.81, 11.01 <sup>a</sup>	14.93, 12.7 <sup>b</sup>	28.55, 8.53 <sup>c</sup>	F = 135.21, df = 2	p < 0.001
<b>Emotion Regulation</b>	64.96, 18.27 <sup>a</sup>	79.13, 20.93 <sup>b</sup>	96.82, 23.59 <sup>c</sup>	F = 151.46, df = 2	p < 0.001
<b>Anxiety</b>	8.82, 9.32 <sup>a</sup>	14.93, 10.29 <sup>b</sup>	23.43, 13.14 <sup>c</sup>	F = 112.5, df = 2	p < 0.001
<b>Depression</b>	7.41, 8.28 <sup>a</sup>	13.66, 10.61 <sup>b</sup>	22.74, 12.72 <sup>c</sup>	F = 133.05, df = 2	p < 0.001
<b>Impulsivity</b>	54.22, 9.52 <sup>a</sup>	59.91, 9.50 <sup>b</sup>	65.60, 11.37 <sup>c</sup>	F = 81.20, df = 2	p < 0.001
<b>Positive Affect</b>	30.47, 9.32 <sup>a</sup>	27.47, 7.70 <sup>b</sup>	23.89, 8.28 <sup>c</sup>	F = 39.25, df = 2	p < 0.001
<b>Negative Affect</b>	11.61, 3.68 <sup>a</sup>	14.05, 5.90 <sup>b</sup>	17.32, 7.51 <sup>c</sup>	F = 60.92, df = 2	p < 0.001
<b>Physical QOL</b>	16.07, 2.65 <sup>a</sup>	14.94, 2.74 <sup>b</sup>	13.58, 3.11 <sup>c</sup>	F = 41.47, df = 2	p < 0.001
<b>Psychol. QOL</b>	15.04, 2.71 <sup>a</sup>	13.24, 2.89 <sup>b</sup>	11.04, 2.90 <sup>c</sup>	F = 102.62, df = 2	p < 0.001
<b>Social QOL</b>	14.83, 3.30 <sup>a</sup>	13.14, 3.69 <sup>b</sup>	11.86, 3.77 <sup>c</sup>	F = 47.51, df = 2	p < 0.001
<b>Environ. QOL</b>	15.26, 2.47 <sup>a</sup>	14.06, 2.61 <sup>b</sup>	12.92, 2.61 <sup>c</sup>	F = 46.74, df = 2	p < 0.001
<b>Neuroticism</b>	27.80, 8.16 <sup>a</sup>	34.69, 7.30 <sup>b</sup>	40.44, 7.65 <sup>c</sup>	F = 717.86, df = 2	p < 0.001
<b>Extraversion</b>	41.59, 8.34 <sup>a</sup>	39.21, 8.46 <sup>b</sup>	37.54, 8.62 <sup>c</sup>	F = 62.43, df = 2	p < 0.001
<b>Openness</b>	40.92, 7.27 <sup>a</sup>	42.80, 7.15 <sup>b</sup>	43.79, 6.96 <sup>c</sup>	F = 45.16, df = 2	p < 0.001
<b>Agreeableness</b>	49.12, 5.92 <sup>a</sup>	47.74, 5.97 <sup>b</sup>	46.12, 6.56 <sup>c</sup>	F = 63.72, df = 2	p < 0.001
<b>Conscientiousness</b>	50.83, 6.12 <sup>a</sup>	47.58, 6.41 <sup>b</sup>	44.04, 6.87 <sup>c</sup>	F = 296.36, df = 2	p < 0.001

*Note.* Post hoc test for ANOVAs is Bonferroni, z-test was used to compare column proportions after Chi-square. Superscript letters that differ across comparison groups indicate statistically significant difference,  $p < .05$ . Any groups that have the same letters did not statistically differ from each other. Scales (see methods for details): BPD symptoms – BSL23, Emotion Regulation – DERS, Anxiety – BAI, Depression – BDI-II, Impulsivity – BIS, Positive and Negative Affect – PANAS, 4 QOL subscales – WHOQOL, Big 5 – IPIP60.

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**



**Figure 1. Self harm** Frequency of participants reporting A) no history of self harm, B) past self-harm, and C) self-harm in the past month is displayed by group: Never BPD (black bars), Remitted BPD (dark gray bars), Current BPD (light gray bars). \* denotes  $p < 0.05$ .

**Draft version 1.0, 5/7/2018. This paper has not been peer reviewed.  
Please do not copy or cite without author's permission.**

**Table S1.** Correlations between measures administered in study phase two.

	BSL	DERS	BAI	BDI	BIS	QOL Health	QOL Psych	QOL SR	QOL Environ	PANAS Pos	PANAS Neg
BSL	-										
DERS	<b>.68</b>	-									
BAI	<b>.75</b>	<b>.54</b>	-								
BDI	<b>.82</b>	<b>.72</b>	<b>.65</b>	-							
BIS	<b>.49</b>	<b>.62</b>	<b>.42</b>	<b>.54</b>	-						
QOL Health	<b>-.54</b>	<b>-.46</b>	<b>-.56</b>	<b>-.61</b>	<b>-.42</b>	-					
QOL Psych.	<b>-.65</b>	<b>-.65</b>	<b>-.51</b>	<b>-.76</b>	<b>-.49</b>	<b>.65</b>	-				
QOL SR	<b>-.39</b>	<b>-.41</b>	<b>-.31</b>	<b>-.51</b>	<b>-.33</b>	<b>.48</b>	<b>.62</b>	-			
QOL Envir.	<b>-.44</b>	<b>-.43</b>	<b>-.40</b>	<b>-.53</b>	<b>-.33</b>	<b>.63</b>	<b>.65</b>	<b>.55</b>	-		
PANAS Pos	<b>-.27</b>	<b>-.42</b>	<b>-.19</b>	<b>-.36</b>	<b>-.32</b>	<b>.28</b>	<b>.56</b>	<b>.40</b>	<b>.34</b>	-	
PANAS Neg	<b>.66</b>	<b>.57</b>	<b>.52</b>	<b>.61</b>	<b>.45</b>	<b>-.39</b>	<b>-.41</b>	<b>-.27</b>	<b>-.32</b>	<b>-.14</b>	-

*Note:* bolded values are statistically significant ( $p < .01$ ).