Trustworthiness Appraisal in Borderline Personality Disorder

A dissertation presented

by

Sara Rose Masland

To

The Department of Psychology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Psychology

Harvard University

Cambridge, Massachusetts

May 2016
Trustworthiness Appraisal in Borderline Personality Disorder

Abstract

Borderline personality (BPD) is a highly impairing illness with marked instability across multiple domains, including affect, interpersonal functioning, identity, and behavior. Within the past 15 years, researchers have sought to understand and characterize deficits in social cognition that might contribute to or arise from affective or interpersonal dysfunction. The purpose of this dissertation is to understand one aspect of impaired social cognition in BPD: biased trust processing. Individuals with borderline features rate others as less trustworthy in laboratory tasks and act accordingly. However, little is known about the influence of affect on ratings or behavior, or whether biased processing is related to real-world functioning or is merely an artifact of the laboratory. The main study included two groups of participants: (1) individuals who met 3 or more diagnostic criteria for borderline personality disorder ($n = 30$) and (2) a control group of individuals who met 2 or fewer diagnostic criteria ($n = 47$).

An experimental paradigm was used to determine whether affectively arousing information has undue influence on trustworthiness ratings for individuals with borderline features. Individuals with borderline features made more negative trustworthiness appraisals overall. Additionally, negative affective information had more impact on trustworthiness ratings for individuals in the borderline features group relative to the control group. This effect was not mediated by rejection sensitivity or moderated by childhood trauma.

A recognition task was used to examine the influence of momentary trustworthiness perceptions on recognition. Overall, individuals who rated faces as less trustworthy during the affective priming task also rated faces as less trustworthy during the recognition task. This
indicates some stability in trustworthiness appraisals. Borderline and control groups did not
differ in recognition accuracy. However, the borderline group expressed significantly less
certainty than the control group about their responses. Incorrect responses in the borderline
group, but not the control group, were also rated as less trustworthy than correct responses.

Social functioning was examined using multiple measures. A pilot study examined the
utility of three new measures of social functioning designed to address perceived flaws in current
measures. A group of community participants (N = 100) completed these measures through
Amazon’s Mechanical Turk, an online crowdsourcing service. Preliminary findings from this
pilot study suggest that these new measures may add meaningfully to our understanding of social
functioning impairments in BPD. These new measures were used alongside already validated
measures to examine how laboratory-demonstrated trustworthiness processing biases relate to
real-world social functioning. Overall, results indicate that despite its reliability and replicability,
the finding that individuals with borderline features have biased trustworthiness appraisal may
not have much practical significance. Trustworthiness bias does not translate to real-world social
functioning, at least in the current study.

Taken together, the findings of the current study show that negative trustworthiness
appraisal is a robust and replicable bias in individuals with borderline personality features.
Consistent with deficits in frontolimbic regulation of affective influence, this bias is augmentable
in the context of negative affective information. However, the bias does not appear to relate to
real-world social functioning. Further work is needed to determine whether laboratory-
demonstrated trust appraisal biases replicate to more ecologically valid scenarios in which
negative trust appraisals may have greater influence on social functioning.
Acknowledgements

Any acknowledgment of support for this dissertation and my intellectual growth must begin with a heartfelt thank you to my advisor, Dr. Jill Hooley. Thank you for artfully supporting my independence while also providing ever-helpful guidance. Thank you for your willingness to help me through blundering sentences and to share in the excitement of my accomplishments. I can only hope to emulate your sage counsel and approachability.

For their time and invaluable guidance, I thank the members of my dissertation committee--Drs. Matt Nock, Jason Mitchell, and Lois Choi-Kain. I thank Dr. Barbara Held, my undergraduate mentor, for her early and continued support and trust in my potential. Additionally, without Dr. John Weisz I would not be where I am today, and I thank him.

For their help with study recruitment, I thank Drs. Karen Jacob and Shauna Dowden. I also thank my lab mates, particularly Tina Chou and Kathryn Fox, for their comraderie and support. Thank you also to Sukriti Drabu, research assistant extraordinaire.

To the fellow members of my clinical science cohort—Chris Rutt, David Dodell-Feder, and Dianne Hezel—putting aside all the joking that has helped keep me smiling over the past six years—I am deeply grateful for your friendship and support. I cannot imagine a better group of people to have experienced graduate school (and life) alongside.

There have been two incredibly supportive people in my home life: my husband, John, and my dog, Crockett. Thank you for tolerating my neglect during busy times and for never failing to make me laugh, even when I stubbornly resist it. You just get it.

This research was supported by the Sackler Foundation, the American Psychological Foundation, and the Department of Psychology at Harvard University.
Dedication

For my brother, who trusted me unfailingly.
# Table of Contents

**CHAPTER 1: INTRODUCTION AND RESEARCH QUESTIONS** ................................................................. 1

**CHAPTER 2: METHOD AND SAMPLE CHARACTERISTICS** ............................................................. 19

- Recruitment ................................................................................................................................. 20
- Procedure ..................................................................................................................................... 21
- Sample Characteristics .................................................................................................................. 32
- Summary ....................................................................................................................................... 43

**CHAPTER 3: AFFECTIVE PRIMING AND TRUSTWORTHINESS APPRAISAL** .............................. 45

- Introduction ................................................................................................................................. 46
- Summary and Predictions ............................................................................................................. 53
- Method .......................................................................................................................................... 54
- Results .......................................................................................................................................... 56
- Discussion ..................................................................................................................................... 74

**CHAPTER 4: FACE RECOGNITION AND TRUSTWORTHINESS APPRAISAL** ........................... 79

- Introduction ................................................................................................................................. 80
- Method .......................................................................................................................................... 85
- Results .......................................................................................................................................... 86
- Discussion ..................................................................................................................................... 95

**CHAPTER 5: MEASURING SOCIAL FUNCTIONING** ................................................................. 99

- Introduction ................................................................................................................................. 100
- New Measure Development ......................................................................................................... 106
- Method .......................................................................................................................................... 109
- Results .......................................................................................................................................... 112
- Discussion ..................................................................................................................................... 122

**CHAPTER 6: TRUSTWORTHINESS APPRAISAL AND SOCIAL FUNCTIONING** ....................... 126

- Introduction ................................................................................................................................. 127
- Method .......................................................................................................................................... 128
- Results .......................................................................................................................................... 129
- Discussion ..................................................................................................................................... 141

**CHAPTER 7: GENERAL DISCUSSION AND CONCLUSIONS** ................................................... 144

- Brief Summary of Primary Results ............................................................................................. 145
- Interpretations and Implications of Findings ............................................................................... 147
- Limitations .................................................................................................................................... 155
- Future Directions .......................................................................................................................... 158
- Conclusion ..................................................................................................................................... 160

**REFERENCES** ............................................................................................................................. 162

**APPENDICES** .............................................................................................................................. 194
Chapter 1: Introduction and Research Questions
Borderline Personality Disorder

Borderline personality disorder (BPD) is fairly common in clinical settings. Prevalence estimates range from 10-15% (Asnaani, Chelminksi, Young, & Zimmerman, 2007; Hyman, 2002) to as high as 22.6% (Korzekwa, Dell, Links, Thababe, & Webb, 2008). In the general population, lifetime prevalence estimates range from 0.7% (Torgersen, Kringlen, & Cramer, 2001) to 5.9% (Grant et al., 2008). The most methodologically rigorous studies converge on a lifetime community prevalence rate in the range of 1-2% (Coid et al., 2009; Lenzenweger et al., 2007). The prevalence of the disorder is particularly problematic when viewed in light of its association with high levels of distress and impairment. Suicide rates are five times higher for individuals with BPD compared to the general population (10%; Paris & Zweig-Frank, 2001).

The disorder is characterized by dysfunction in three dimensions: affective, behavioral, and interpersonal (Skodol et al., 2002b). Dysfunction in these three areas often manifests as instability, and many have now come to understand the disorder as one of “stable instability” (see Hooley, Cole, & Gironde, 2012). Indeed, the World Health Organization calls BPD “emotionally unstable disorder” in its International Classification of Diseases (ICD-10; WHO, 1992). The idea of instability is further reflected in the DSM-5 description of BPD as involving “a pervasive pattern of instability of interpersonal relationships, self-image, and affects, and marked impulsivity” (DSM-5; APA, 2013, p. 663).

Instability in interpersonal relationships and processes manifests as alternating idealization and devaluation of relationships, fears of abandonment, and inappropriate expression of anger (APA, 2013). Interpersonal instability has been a focus of clinical theories of BPD, and many consider it to be the central dysfunction in the disorder (e.g., Gunderson, 1996; Kernberg, 1967). Indeed, the affective and behavioral dysfunction observed in BPD often results from interpersonal difficulty. Aspects of interpersonal vulnerability have been shown to mediate the
relationship between borderline features and emotional reactivity (Dixon-Gordon, Yiu, & Chapman, 2012). Psychological distress and negative affect in BPD are most commonly caused by interpersonal stressors, including abandonment or rejection (Stiglmayr et al., 2005). Compared to both individuals with other personality disorders and those with no personality pathology, daily interaction is marked by greater negative affect for individuals with BPD (Stepp, Pilkonis, Yaggi, Morse, & Feske, 2009). Stimuli evoking themes of rejection and abandonment, decidedly interpersonal events, are more emotionally and physiologically arousing for individuals with BPD compared to healthy controls (Limberg et al., 2011). Additionally, real or imagined interpersonal conflict (Levy, 2005) or interpersonal events (Brodsky et al., 2006) often trigger self-destructive behavior, including self-harm and suicidality for individuals with BPD. Although one might expect that aggression (behavioral dysfunction) might cause interpersonal dysfunction, evidence suggests that it is more often the other way around (Herr et al., 2013). For example, a bias to perceive rejection in interpersonal interactions may lead an individual with BPD to act aggressively.

**Borderline Personality Disorder and Interpersonal Dysfunction**

Unsurprisingly, interpersonal instability results in significant interpersonal dysfunction in BPD. Interpersonal dysfunction may be characterized as a primary phenotype of BPD (Gunderson, 2007). Empirical evidence has shown that BPD symptoms are strongly related to interpersonal sensitivity, aggression (Bouchard et al., 2009; Lejuez et al., 2003), ambivalence, need for approval, and lack of sociability (Stepp et al., 2012), all of which may contribute to interpersonal dysfunction. Additionally, individuals with BPD commonly have fearful, ambivalent, or preoccupied attachment styles (Agrawal et al., 2004; Choi-Kain et al., 2009) are hypersensitive to rejection (Staebler, Helbing, Rosenbach, & Renneberg, 2011), and may have
expectations of extreme inclusion in social interactions (De Panfilis et al., 2015). In one study, self-reported interpersonal problems among those with BPD clustered into vindictiveness, submissiveness, nonassertion, exploitability, and social avoidance (Salzer et al., 2013).

The magnitude of impairments in interpersonal functioning is greater in BPD (and also in schizotypal personality disorder; SPD) than it is in other personality disorders (Skodol et al., 2002a). Also noteworthy is that interpersonal dysfunction in nonclinical samples of people with borderline features is nearly as severe as it is in clinical samples (e.g., Tolpin et al., 2004). Furthermore, impairment in social relationships is very stable for personality disorders in general (Skodol et al., 2005) and for BPD in particular, with little fluctuation in severity of dysfunction (Wright et al., 2013). Importantly, interpersonal dysfunction in BPD endures even during symptom remission. In a 10-year longitudinal study of BPD course, Gunderson and colleagues (2011) found that 85% of cases remitted. Despite this, social function improved only moderately and remained severely impaired, more so for individuals with BPD than for those with major depressive disorder or cluster C personality disorders (Gunderson et al., 2011). Similarly, Trull and colleagues (1997) found that young adults with sub-clinical borderline features were more likely to have interpersonal problems at a 2-year follow up than were their peers without borderline features. Thus interpersonal problems persist over time even in the absence of full diagnosis and also in the face of symptom remission. Findings such as these highlight the need to better identify mechanisms that might perpetuate interpersonal dysfunction in BPD.

Interpersonal dysfunction manifests problematically in therapeutic relationships, and is frequently identified by clinicians as a barrier to successful psychotherapy. Troubled interpersonal functioning also manifests in general social networks, romantic relationships, and family relationships. Some evidence suggests that BPD is more closely associated with
dysfunction in romantic relationships, interactions with friends, and in work environments (Hill et al., 2008). Generally, individuals with borderline personality tend to have tumultuous, stormy relationships that are short in duration and cause a great degree of emotional distress and pain to the individual. Patients with BPD express intense needs for closeness and a paradoxically equally intense fear of rejection or abandonment (Gunderson & Ruth-Lyons, 2008).

**Interpersonal dysfunction in general social networks.** BPD is associated with friendships of shorter overall length, lack of close confidants and fewer social activities as adolescents, and the experience of decreased reciprocity within social networks (Daley, Burge, & Hammen, 2000; Stepp et al., 2009). Strikingly, despite clear disruptions in general social functioning, people with BPD tend to report internal feelings of connectedness, wellbeing and stability during interpersonal contact (Stanley & Siever, 2010). In addition, social participation plays an important role in alleviating their feelings of sadness (Renneberg et al., 2012).

**Interpersonal dysfunction in romantic relationships.** Empirical research investigating romantic functioning in borderline personality disorder reveals significant levels of dysfunction, particularly among women. Daley and colleagues (2000) examined the number and quality of relationships in young women diagnosed with BPD in a longitudinal study. They showed that the number and severity of borderline symptoms was positively associated with frequency of conflict in intimate relationships and a greater risk of partner abuse. In addition, a diagnosis of BPD was associated with a lower likelihood of being married and a higher number of romantic breakups. Other work has shown that people with borderline features are also less likely to be married (Swartz et al., 1990), experience greater conflict in romantic relationships (Chen et al., 2004), display more negative skills during problem-solving tasks with their partners (Lavner, Lamkin,
& Miller, 2015), report lower satisfaction with relationships (Bouchard et al., 2009), and report
greater distress earlier in the course of relationships (Lavner et al., 2015).

**Interpersonal dysfunction in family relationships.** Research on family relationships
among patients diagnosed with BPD tends to focus on the role of the family in the etiology of
BPD. This work has identified that patterns of early attachment and adverse childhood
experiences are particularly important to understanding BPD. Many studies support the idea that
individuals with BPD tend to have more difficult and painful early family experiences
(Gunderson & Ruth-Lyon, 2008). Similarly, people with BPD are more frequently victims of
sexual assault by a known assailant (Preißler, Dzlobek, Ritter, Heekeren & Roepke, 2010).
Furthermore, the severity of assault or mistreatment is linked to later observed severity of
borderline symptoms (Lieb et al., 2004). A longitudinal study revealed that children with
maternal withdrawal or separation during the first five years of life demonstrated significantly
higher BPD symptom elevations from early adolescence into middle adulthood (Steele & Siever,
2010). However, it is important to note that many individuals with BPD do not report significant
abuse, maltreatment, or attachment disruption. More subtle processes may be at play, including
poor “fit” between an emotionally vulnerable child and a parent who has difficulty validating
emotional experiences (Linehan, 1993).

**Processes Underlying Interpersonal Dysfunction**

Interpersonal dysfunction in BPD is significant, enduring, and generalized, affecting
relationships of all types. Moreover, as a result of research conducted to date, the phenotypic
expression of interpersonal dysfunction in BPD is also now fairly well characterized. Yet the
processes that support this interpersonal dysfunction are much less clear. Two key psychological
endophenotypes may be important here. These are failures in mentalization and rejection sensitivity (Gunderson, 2007).

**Mentalization.** Mentalization is the process of understanding mental states (e.g., beliefs, intentions, feelings) in oneself and others. Disturbed mentalization may develop via a number of potential pathways in BPD and leads to interpersonal hypersensitivity, emotional contagion, and poor understanding of affective and cognitive states and how they interact (see Fonagy, Luyten, & Strathearn, 2011). Theories of disturbed mentalization in BPD are similar to ideas offered in early work by Kernberg (1975, 1984). Kernberg has maintained that perceptual disturbances in BPD arise from undifferentiated self/other representations and the use of defense mechanisms (e.g., splitting, projective identification) in the absence of problems with reality testing. In both the case of mentalization theory and Kernberg’s more classically psychoanalytic theory, the self is improperly differentiated from others, leading to poor understanding of one’s own mental states as well as the mental states of others. This in turn gives rise to a number of biases that manifest problematically in interpersonal relationships. Recent neuroimaging evidence from Beeney and colleagues (in press) suggests that individuals with BPD have less integrated and more negative self-representations and may “hypermentalize.” These individuals showed greater more activation in the theory of mind network and less activation in areas of visual, sensory, and motor cortex than controls when considering representations of close others. The authors take this as an indication that individuals with BPD make greater attempts to interpret the intentions and beliefs of others, but rely less on important external cues for this understanding. This is congruent with other evidence suggesting that individuals with BPD are less able to use information in the environment to determine how others are perceiving them; instead they rely more on their own perceptions (Carlson & Oltmanns, 2015).
**Social cognition in borderline personality disorder.** Consistent with the notion that individuals with BPD are more sensitive to rejection and have more difficulty understanding the mental states of others, a number of disturbances in social cognition have been observed. For example, individuals with BPD show a bias in social interactions that manifests in negative emotional states. BPD features predict greater negative impact and feelings of emotional loss when romantic partners initiate both positive and negative experiences, owing to negative appraisals of those experiences (Bhatia et al., 2013).

The tendency to appraise social interactions more negatively may be due in part to disturbances in facial emotion recognition. A great number of studies have now investigated whether facial emotion recognition is impaired in BPD. Unfortunately, results have been highly contradictory perhaps in part due to methodological variance and differences in sample characteristics (see Domes et al., 2009; Mitchell, Dickens, & Picchioni, 2014). Some have concluded that BPD is associated with poorer recognition of facial emotion (e.g., Bland et al., 2004; Unoka et al., 2011), while others have found no differences between people with BPD and healthy controls (e.g., Dyck et al., 2009; Minzenberg, 2006). Still others have found that people with BPD are better able to identify facial emotions (e.g., Wagner & Linehan, 1999). A recent meta-analysis supports the idea of a fairly consistent negative response bias for neutral and ambiguous faces in BPD, as well as no differences between healthy controls and individuals with BPD in the ability to identify negative emotions (Mitchell et al., 2014). However, evidence published by Veague and Hooley (2014) after this meta-analysis suggests that women with BPD, irrespective of abuse history, detect anger more quickly and are more likely to misidentify anger in male faces. They are also slower to recognize happiness in male faces. Evidence from psychophysiological work suggests that individuals with BPD have a reduced threshold for
perceiving facial anger driven by very early perceptual alterations (Hidalgo et al., 2015).

Although further work is needed to draw firm conclusions about emotion recognition in BPD, disturbances in emotion recognition are one plausible mechanism by which individuals with BPD may come to view social interactions more negatively. Other plausible mechanisms may be found in other aspects of social cognition, including trait appraisal

**Trait appraisal in borderline personality disorder.** In recent years, work has shifted away from emotion recognition and toward investigation of trait appraisal in BPD. Implicit trait appraisal is a primary shortcut during which we make quick, complex judgments about the traits or personality of others. Based entirely on facial characteristics we make enduring judgments about a range of traits, including dominance, likability, and competence (Todorov et al., 2005). Unlike emotion recognition, trait appraisal is the process of ascribing personality characteristics—rather than emotional states—to an individual. Trait appraisals are made in as little as 100ms upon viewing a face; longer exposure to faces does little to change initial appraisals but increases confidence in initial judgments (Willis & Todorov, 2006). These appraisals are important to interpersonal functioning as they serve as heuristic shortcuts by which individuals make inferences about the mental states of others and predict their future behavior. Judging a person as untrustworthy, for example, allows us to make predictions about betrayal and to infer that the person might be thinking about how to swindle us.

Trait judgments have important behavioral consequences. For example, initial, rapidly-made ratings of competence predicted voting behavior in a congressional election, indicating that this quick online judgment had as great or greater an influence on behavior than rational decision-making (Todorov et al., 2005). Recently, researchers have begun to evaluate the process of trust or trustworthiness appraisal in BPD, as clinicians and scientists have long taken note of
how difficult it is for individuals with BPD to trust others. Moreover, single interactions and appraisals form the foundations of more enduring, and infinitely more complex relationships. Understanding these early and implicit appraisals may offer insights into how more complex relationships unfold and how this process may be disturbed for individuals with BPD.

This process has been examined in BPD using two types of paradigms: visual appraisal paradigms and trust games. Visual appraisal paradigms require the individual to view pictures of faces or videos and to judge, based only on limited information, how trustworthy the person is. Fertuck, Grinband, and Stanley (2013) showed individuals a series of faces that varied on either a fear or trust dimension, and asked them to rate the trustworthiness of each face. Compared to controls, the BPD group rated faces as less trustworthy. The authors attribute this to a response bias rather than heightened sensitivity or discriminability in the BPD group. Importantly, they also found that this trust bias occurred independently of fear appraisal, indicating that less trustworthy faces are not simply those that evoke greater fear response. Nicol et al. (2013) found that relative to healthy controls, individuals with BPD rate faces as less trustworthy and less approachable. Approachability ratings were significantly correlated with scores on the Childhood Trauma Questionnaire (CTQ; Berenstein & Fink, 1997) within the BPD group, but trustworthiness ratings were not. Others have shown that trustworthiness appraisals in non-clinical populations do not vary as a function of true differences in how much an individual should be trusted based on actual behavior (Rule et al., 2013). Using a similar paradigm to Fertuck et al. (2013) with both facial and limited behavioral information available to subjects, Arntz and Veen (2001) found that individuals with BPD rate others more negatively in general when viewing short movie clips. Relatedly, Barnow et al. (2009) found that such individuals rate others as more aggressive than do either healthy controls or depressed patients.
Some have also attempted to understand the interpersonal dysfunction observed in BPD by examining how individuals who score high on constructs related to the disorder, but who do not have the disorder, process social information. For example, Miano et al. (2013) found that individuals with sub-clinical BPD features also make more negative trust appraisals of neutral faces, and that this effect is moderated by rejection sensitivity. Individuals with BPD and those high on borderline traits form negative, untrusting initial impressions of others.

Trustworthiness/trust appraisal bias does not necessarily imply behaviors consistent with bias. For example, it is possible in principle that individuals with BPD may rate others as less trustworthy, but may act as though others are indeed trustworthy. This is likely not the case. To examine the behavioral consequences of negative impression formation, researchers have used monetary investment paradigms. These paradigms are presented as games in which individuals with BPD act either as investors or trustees. Investors choose whether to invest money in trustees, who can then cooperate by returning money to the investor. The investor and trustee are able to earn more money together through cooperative exchange than they would through conservative investment. However, cooperative exchange includes risk. Investing in the other party may backfire if the other party chooses not to cooperate by sending money back. These monetary exchanges thus require mutual trust for optimal performance. Participants must, therefore, form initial impressions of others’ trustworthiness and then cooperate with those others toward common monetary ends.

In contrast to healthy control participants, when people with BPD play these games, they are more concerned with perceived fairness/trustworthiness than facial cues of happiness, neutrality, or anger from confederates, even when the emotional cues are consistently linked with trustee behavior (Franzen et al., 2011). In other words, trustworthiness appraisals dictate their
behavior more than emotional appraisals—individuals with BPD are less likely to rely on cues of anger than cues of trustworthiness when deciding whether to invest in another. Unoka et al. (2009) used a monetary investment paradigm in which individuals with BPD acted as investors. Compared to control and depressed groups, the BPD group made smaller investments overall during the course of five trials, indicating less trust in the trustee. This group also reported that they were less optimistic about the outcome of the investment game prior to beginning. Similar findings are observed when roles are reversed. Participants with BPD, acting as trustees rather than as investors, rupture cooperation and trust by sending less money back to the investor (King-Casas et al., 2008). This indicates that individuals with BPD are both less trusting and less cooperative in these types of economic games.

Evidence from paradigms focused on visual appraisal and trusting behavior in economic exchanges converges to show that trust processing biases exist and may have important negative behavioral consequences for individuals with borderline pathology. However, empirical investigations of trust processing in BPD have thus far focused entirely on well-controlled laboratory paradigms. Little is known about contextual factors (e.g., emotion) that may influence trustworthiness appraisal or what mechanisms underlie or influence aberrant trust processing. Furthermore, work to date has not sought to link trust processing to real-world interpersonal/social functioning explicitly. Approaching the world as though others are less trustworthy may contribute to the interpersonal dysfunction characteristic to BPD. Better understanding of biased trust processing may provide insights into how we might seek to improve interpersonal functioning for those with this disorder. In this dissertation I evaluated potential mechanisms underlying trust appraisal and sought to determine whether trust processing biases relate to real-world functioning. In Chapter 3, I evaluate the influence of
affective information on trustworthiness appraisals. In Chapter 5, I consider issues related to measuring social functioning in BPD, and in Chapter 6 I test whether laboratory-observed trust processing biases are associated with reports of social functioning.

**Affective Instability and Trust Appraisal**

Affective dysfunction, like interpersonal dysfunction, is a characteristic of BPD. The disorder is marked by affective instability. As has already been discussed, aversive interpersonal experiences often result in dysregulated emotion for people with borderline pathology. But this is not a one-way street. Affective instability may also contribute directly to interpersonal dysfunction. This is well-illustrated in the diagnostic criteria for the disorder. For example, inappropriate expressions of anger, one symptom of BPD, may result from an inability to control strong emotions. Inappropriate displays of anger might also be expected to cause ruptures in interpersonal relationships.

Some argue that affective instability, rather than interpersonal dysfunction, is the core of borderline pathology. Linehan (1993), for example, argues that individuals who go on to develop BPD are biologically predisposed to emotional vulnerability, and that this diathesis lies at the root of the disorder. Indeed, some evidence suggests that affective instability can drive behavioral and interpersonal dysfunction (Tragesser et al., 2007) and that affective instability, along with anger, is one of the two most stable characteristics of BPD (McGlashan et al., 2005). Emotional dysregulation may also maintain enduring borderline pathology (Stepp et al., 2013). Individuals with BPD have higher emotional intensity, more chronic negative affect, and higher baseline arousal than healthy controls (Scott et al., 2013). They also show hyperarousal when processing social or emotional material compared to controls (Ruocco et al., 2012a).
These findings are not surprising, given that many individuals with BPD report many early life stressors and high rates of abuse (e.g., Zanarini, 2000). These are factors that are known to dysregulate the hypothalamic-pituitary-adrenal (HPA) axis in individuals with BPD (e.g., Rinne et al., 2002). HPA dysregulation may render individuals more vulnerable to stress. Evidence suggests that the HPA axis is hyper-reactive in BPD, and that the interaction of the resulting hypercortisolism with a disturbed serotonin system may result in affective instability, including greater responses to stressful events (Wingenfeld et al., 2000). Notably, social evaluative stress, which is decidedly interpersonal in nature, elicits the greatest cortisol responses from the HPA system, particularly when social evaluative stress is considered uncontrollable (Dickerson & Kemeny, 2004). Dysregulated HPA axis activity may interact with other physiological vulnerabilities to render individuals with BPD particularly prone to emotional lability. For example, emerging evidence suggests that the amygdala is hyperactive in BPD patients relative to healthy controls when viewing faces, regardless of facial expression (Herpertz et al., 2001; Donegan et al., 2003).

Although it is now understood that both interpersonal dysfunction and affective instability are central to BPD, no work to date has examined how affective arousal might influence trust appraisal. Understanding the influence of affective arousal on trait appraisal, particularly trust appraisal, is paramount for our understanding of emotional lability and interpersonal difficulty in BPD. In Chapter 3, I examined the effect of affective information on trust appraisals for individuals with borderline pathology compared to a control group.

**Trust Judgments and Memory for Faces**

Individuals with BPD perform more poorly than healthy controls across several domains of neuropsychological functioning including memory. However, the average effect sizes of these
differences are small and likely not clinically useful (Ruocco, 2005). It may be the case that cold cognition, as assessed with tasks using neutral stimuli in the absence of emotionally relevant information, differs only slightly for individuals with BPD. Cognitive processing difficulties, particularly those related to memory, may be particularly prominent when information is affectively salient. Indeed, individuals with BPD show memory dysfunction relative to healthy controls in the presence of negative emotional information (Mensebach et al., 2009) and have difficulty with directed forgetting of this affectively salient information (Korfine & Hooley, 2000). This means that they are less able to exert control over this information—it is more likely to persist in memory even with conscious efforts to avoid encoding and retrieval.

Interpersonal events are more emotionally salient for individuals with borderline pathology, and as discussed, these individuals are more reactive to these events. Rating the trustworthiness of others is a necessarily interpersonal process, and we might expect that faces that are viewed as less trustworthy are more affectively salient. Low trustworthiness may signal danger (e.g., betrayal, rejection) in an affectively loaded manner. The perception of a face as untrustworthy may thus interfere with retrieval processes when an individual with borderline features is asked to remember whether he has seen the face before. The influence of trustworthiness appraisals on memory for faces has not yet been tested in BPD. Here, in Chapter 4, I sought to determine whether the perception of low trustworthiness does indeed impair memory for previously viewed faces.

**Trust Processing and Social Functioning**

Trust processing biases have been demonstrated in a number of laboratory experiments and may contribute to interpersonal functioning outside of the laboratory. However, to date, no studies of trust processing have attempted to relate biases found in the laboratory directly to real-
world social functioning. Therefore we cannot be sure if the biases we observe in the laboratory are of pragmatic significance. This is an important omission in the literature if we are to capitalize upon our understanding of trust processing to improve treatment for BPD.

Important obstacles exist in relating trust processing biases to real-world social function. These include (1) the difficulty of measuring actual social functioning in a disorder marked by negative judgments even of positive interpersonal interactions (e.g., Bhatia et al., 2013); and (2) the lack of a measure that, on the item level, clearly taps into how well a person is functioning overall in relationships and other social contexts. One commonly used measure of social function in BPD is the Inventory of Interpersonal Problems Personality Disorders 25 (IIP-PD-25; Kim & Pilkonis, 1999). While this measure has strong psychometric properties, examining the items (e.g., “I am too envious of other people”) reveals that they may in some cases be a better measure of negative self-concept or general negativity than actual functioning. Other scales are limited to particular areas of functioning. Some studies use the interpersonal stability criterion of BPD as their measure of social function, which implicitly assumes that individuals who have borderline pathology but who not meet this criterion are functioning well. Here, in Chapter 5, I developed and piloted a more comprehensive measure of social functioning. I then used this measure alongside previously validated measures in Chapter 6 to determine how trust processing biases observed in the laboratory relate to real-world social functioning.

**Current Research**

Converging evidence now compellingly shows that individuals with borderline personality disorder and subthreshold borderline personality features rate others as less trustworthy and act accordingly. Biased trustworthiness appraisals may support troubled interpersonal functioning for individuals with borderline pathology. However, to the author’s
best knowledge, no work to date has examined whether trustworthiness bias measured in laboratory settings is associated with real-world social functioning. Social/interpersonal dysfunction is a core feature of borderline pathology. Affective lability is another core feature of borderline pathology, and the interplay of affective lability with the social cognitive process of trustworthiness appraisal is poorly understood. The current research examines this intersection by considering the influence of affective information on trustworthiness appraisals. Additionally, trustworthiness appraisals may be influenced by affective context, but they may also denote affective information in and of themselves. They may thus have consequences for other cognitive processes, including memory. The current research examines the possibility that perceptions of trustworthiness may influence a borderline individual’s ability to recall whether they have previously perceived or encountered the person whom they are rating.

To examine these issues, participants were recruited from the general community on the basis of borderline pathology and general social functioning. Each participant was placed in one of two groups delineated by number of borderline personality traits endorsed. Individuals who endorsed fewer than 3 borderline traits were categorized as control participants. Individuals who endorsed three or more borderline traits were placed into a “borderline features” group. Study participants completed several psychometric assessment measures at home and in the laboratory, a diagnostic assessment, detailed measures of social functioning, and two behavioral tasks. Behavioral tasks included (1) an affective priming task designed to examine the influence of affective information on trustworthiness appraisals; and (2) a memory task during which participants were asked to remember whether they had seen each of a series of faces in the earlier task, to rate how certain they were of their memory judgments, and to rate the trustworthiness of each face.
Participant recruitment and general sample characteristics are discussed in greater detail in Chapter 2. Chapter 3 reports the results of an affective priming task and considers how affective priming may have differential impact on individuals with borderline features compared to controls. Results from the second behavioral task, during which participants were asked to remember whether they had seen each in a series of faces before, are presented in Chapter 4. Chapter 5 critically considers common methods for assessing social functioning in BPD and reports pilot results pertaining to a new measure of social functioning. This new measure is included alongside previously validated measures in Chapter 6 to determine whether laboratory-observed trustworthiness appraisal biases are associated with social functioning. Finally, in Chapter 7, I provide a general summary of the findings and offer concluding comments.
Chapter 2: Method and Sample Characteristics
The overall purpose of this dissertation was to better understand biased trust processing in BPD. Individuals with borderline features rate others as less trustworthy in laboratory tasks and act accordingly. However, little is known about the influence of affect on trustworthiness ratings or trusting behavior, or whether biased processing is related to real-world functioning.

Here I describe the recruitment process for the main study participants and present data on the demographic and clinical characteristics of the sample. I also describe laboratory procedures and measures for the main portion of the current study. The main study included two groups of participants: (1) individuals who met 3 or more diagnostic criteria for borderline personality disorder \( n = 30 \) and (2) a control group of individuals who met 2 or fewer diagnostic criteria for borderline personality disorder \( n = 47 \). Pilot testing of new social functioning measures with a separate sample from the one described here is detailed in Chapter 5.

**Recruitment**

Adult participants were recruited from the greater Boston area using three primary methods. The first method used flyers and Craigslist advertisements to draw participants directly from the community. Control participants were required to “have no history of mental illness,” or to consider themselves “healthy individuals with strong social relationships.” Recruitment materials for individuals with borderline personality traits included such questions as “Has anyone ever told you that you have borderline personality?”, “Do you ever feel empty? Do you often feel bored? Or do you have trouble being alone?”, and “Do your moods change often? Do you have trouble keeping relationships with others?”

The second recruitment approach took advantage of the fact that, at the time of participant recruitment, another study in the same laboratory was recruiting individuals with a history of self-harm. If participants in this study indicated that they were interested in
participating in other studies, and if their questionnaire responses indicated the likely presence of borderline traits (i.e., 3 or more borderline traits endorsed on the SNAP-2, described below), they were invited to participate in the present study. A cut-off of 3 or more BPD traits is slightly lower than the SNAP’s cutoff for “borderline features” (four criteria). However, it is in-line with one of the largest studies of BPD currently ongoing: the Collaborative Longitudinal Personality Disorder Study (CLPS; see Gunderson et al., 2011). This study considers BPD in remission when 2 or fewer criteria are met. Thus those individuals who meet 3 or more criteria do not qualify as in remission and are considered to have significant levels of borderline pathology.

The third recruitment approach involved requesting the assistance of several community providers of mental health care. Two providers allowed me to attend an outpatient therapy group for individuals who had completed residential treatment for borderline personality disorder at McLean Hospital’s Gunderson Residence. This outpatient group was not directly affiliated with McLean Hospital, although the clinicians who run the group are employed at McLean. I presented the study to the group and provided patients with flyers directly. Additionally, another outpatient psychotherapist who specializes in treating borderline personality, and with whom I have worked in the past, directly provided study flyers to a number of her patients.

Exclusion criteria included clinically significant vision or hearing problems not corrected to normal by eyeglasses or hearing aids, history of significant head trauma (with loss of consciousness), neurological illness (e.g., multiple sclerosis), current use of antipsychotic medication, and diagnosis of a psychotic disorder.

Procedure

Participants were first screened via phone to ensure that they did not meet any of the exclusion criteria (see above). Pre-screening did not include any assessment of borderline
personality, as I sought to enroll participants with a full range (0-9) of borderline traits. I monitored borderline trait endorsement as participants completed the study to ensure recruitment was successfully targeting individuals endorsing at least 3 or 4 of these traits. Once screened, participants were provided with an ID number over the phone and a link via email. This link connected participants with an initial consent form and a set of questionnaires to be completed via Qualtrics before visiting the laboratory. Completion of these questionnaires required approximately 1 hour of time. Participants were asked to complete these measures during the week before their laboratory visits. Completion of these measures at home shortened the time required during the laboratory visit, thereby decreasing the possibility that participants might become fatigued or disinterested in the study.

When participants reported to the lab, I first provided them with more information about the study and reviewed a second consent form with them. Immediately following the informed consent process, participants completed the SNAP-2, the BDI-II, and the CTQ (see below). Once these self-report measures were completed, each participant participated in an interview that included diagnostic measures (the MINI and relevant sections of the SCID-II; see below), and the WTAR, a brief measure of premorbid intellectual functioning (see below). Each participant was then offered a break and snacks before he or she completed the behavioral trust task with measures of mood before and after (see Chapter 3). Following the behavioral trust task, participants completed measures of social functioning for a period of 12-15 minutes before completing a memory task (see Chapter 4). Following completion of study measures participants were debriefed and provided with payment ($10/hour in cash). Lab procedures required 2-3 hours for each participant in addition to the 1 hour spent on questionnaires at home.
Response to advertisements (craigslist, flyers, community clinicians) → Phone Screening (general exclusion criteria assessed) → Questionnaires sent and completed via Qualtrics → Laboratory Session

Figure 2.1. General recruitment flowchart

Study Explanation and Consent → Self-Report Measures → Diagnostic Assessment → Snack Break

Trust Appraisal Task (with Mood Measures Pre/Post) → Delay (Distractor Questionnaire) → Memory/Recognition Task → Debriefing

Figure 2.2. Laboratory procedures flowchart
Measures

Measures included a diagnostic assessment, self-report questionnaires, and behavioral paradigms. Behavioral tasks and measures relevant to the behavioral tasks (i.e., the trust task and the mood measurements given before and after the trust task, the memory task, and the anchoring/adjustment task) are described in the following chapters. Social functioning measures are also described in greater detail in relevant chapters.

Diagnostic Assessment

Participants completed two diagnostic assessments. First, they completed the Mini-International Neuropsychiatric Interview 5.0 (MINI; Sheehan et al., 1998). This is a short structured diagnostic interview designed to query for the presence of DSM-IV Axis I disorders. The MINI is commonly used in epidemiological studies and clinical trials owing to its short administration time (approximately 15 minutes) and low training burden. The reliability and validity of the MINI have been studied extensively in comparison to other structured interviewed with established validity. The results broadly indicate that the MINI is a reliable and valid tool (e.g., Sheehan et al., 1997; Lecrubier et al., 1997).

A DSM-5 version of the MINI was not available prior to study enrollment. However, relevant modifications were made to assess for DSM-5 disorders (e.g., the requirement that an individual react with “fear, helplessness, or horror” during a traumatic event was removed from the diagnostic algorithm for posttraumatic stress disorder).

Second, participants completed two sections of the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II; First et al., 1997). These sections included questions relevant to the assessment of borderline and schizotypal personality disorder. The section on borderline personality was included to directly assess borderline traits via interview. A measure of schizotypal personality was included because personality disorders are generally impairing,
particularly with regard to social functioning. Including the measurement of a second personality disorder within the same sample was intended to allow for claims about the specificity of social functioning effects to borderline personality, and to consider what effects may be due to borderline personality versus personality disorder in general.

The SCID-II borderline and schizotypal personality sections have high reliability in terms of inter-rater agreement (Kappa = .91 for borderline personality, .91 for schizotypal personality; Maffei et al., 1997); and trait intraclass correlations (.93 for borderline personality, .62 for schizotypal personality; Lobbestael, Leurgans, & Arntz, 2011). Examinations of categorical and dimensional inter-rater reliability confirm that agreement is high (e.g., Lobbestael et al., 2011). Other work has shown that individuals diagnosed with borderline personality using the SCID-II have higher risk of comorbidity and lower occupational status than individuals without personality disorder (Andion et al., 2012). Whereas some work has challenged the validity of the SCID-II and our current model of personality disorder more broadly, other work has shown that the SCID-II performs best in terms of reliability and validity when diagnosing borderline personality. The borderline subsection has satisfactory convergent validity, divergent validity, association with general personality, and relation to functional impairment. The same is true for the schizotypal personality section, with the exception that schizotypy on the SCID-II does not show relation to any general personality traits (i.e., the Big 5; Ryder, Costa, & Bagby, 2007).

**Intellectual Ability**

IQ estimates were derived using the Wechsler Test of Adult Reading (WTAR; Holdnack, 2001). The WTAR takes fewer than 5 minutes to administer and is co-normed with the Wechsler Adult Intelligence and Memory Scales (WAIS-III, WMS-III). Raw scores on the WTAR were
transformed to estimated WAIS-III IQ scores using normative data corrected for age, education, sex, and ethnicity, as appropriate.

**Self-Report of Borderline Personality**

Although BPD traits were captured during diagnostic interviewing as well as on the SNAP-2 (described below), we also included relevant items from the Personality Assessment Inventory (PAI; Morey, 1991) to capture dimensions of borderline pathology. The full PAI has 344 items that are answered on a 4-point scale (where 1 means “not true at all” and 4 means “very true”). The full PAI measures both personality and psychopathology and includes 4 validity scales, 11 clinical scales, 5 treatment scales, and 2 interpersonal scales. The measure was developed using a construct validation framework with rational and quantitative methods.

Only the 24 items comprising the borderline features scale of the PAI (PAI-BOR) were included. The borderline scale has four subscales: affective instability (BOR-A), identity problems (BOR-I), negative relationships (BOR-N), and self-harm (BOR-S). Scores on the BOR scale above a T score of 90 indicate the likely presence of borderline personality disorder (Morey, 2003). In some studies, lower T scores have proven to blend specificity and sensitivity: a T score greater than or equal to 70 accurately captured 82% of borderline diagnoses in one study (Bell-Pringle, Pate, & Brown, 1997), while this same T score accurately classified borderline personality at a rate of 73% in other research (Stein, Pinkser-Aspen, & Hilsenroth, 2007). Others have shown that a raw score greater than or equal to 38 captures clinically significant borderline symptoms (Trull, 1995). Notably, Morey (2003) writes that individuals with high BOR scores “tend to be hypersensitive to any signs of rejection and may rapidly attempt to test the therapist’s commitment or trustworthiness” (p. 114).
Self-Report of Schizotypy

All participants completed the Schizotypal Personality Questionnaire (SPQ; Raine, 1991). The SPQ is a 74-item scale developed based on DSM-III-R criteria for schizotypal personality disorder. The scale has high internal consistency ($\alpha = .91$), test-retest reliability over a two-month interval (.82), convergent and discriminant validity, and criterion validity (Raine, 1991). Items factor into nine subscales, including ideas of reference, excessive social anxiety, odd beliefs or magical thinking, odd or eccentric behavior, unusual perceptual experiences, lack of close friends, odd speech, constricted affect, and suspiciousness.

Schedule for Nonadaptive and Adaptive Personality (SNAP-2; Clark, 2006).

The Schedule for Nonadaptive and Adaptive Personality-2 (SNAP-2; Clark, 2006), is a 390-item true/false self-report questionnaire. The measure is factor analytically derived and includes 12 trait scales (e.g., mistrust, manipulativeness, aggression) 3 temperament scales (negative temperament, positive temperament, and disinhibition), 12 diagnostic scales (e.g., borderline, paranoid, narcissistic), 6 validity scales (e.g., variable response inconsistency, desirable response inconsistency), and one overall invalidity scale to identify invalid profiles. Of particular note, the SNAP-2 has subscales for borderline traits and general mistrust.

The psychometric properties of the SNAP-2 have been evaluated in diverse populations, including normative adults, college students, mixed patients, back pain patients, and patients with recurrent depression. Internal consistencies ($\alpha$) for the trait and temperament subscales average .80 to .84 and range from .63 to .93 across these populations. Test-retest correlations over 1-2 months in college samples average .80; over 7 days to 4.5 months in community adults they average .87; and over a shorter term they average .81 in patient populations (Ro, Stringer, & Clark, 2012). Internal consistencies ($\alpha$) of the diagnostic subscales range from .65 to .89 for
normative and patient samples and show acceptable stability with test-retest reliabilities ranging from .77 to .91 (Clark, 2006). Women consistently score higher on negative temperament and lower on disinhibition and manipulativeness than men (Ro, Stringer, & Clark, 2012).

**Depression**

Depressive symptoms were assessed using the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996). The BDI-II is a 21-item multiple-choice self-report scale that assesses the severity of depressive symptoms. The scale has strong internal consistency (α = .91) and high test-retest reliability (.93) over a 1-week period.

**Anxiety: Trait and Social**

Willis, Dodd, and Polermo (2013) have reported that higher levels of trait anxiety predict more negative ratings of trustworthiness. Accordingly, we measured trait anxiety using the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger, 1983). The scale has satisfactory internal consistency (α ranges from .87 to .93) and test-retest reliability (ranging from .73 to .86 after 3-15 weeks; Spielberger & Reheiser, 2009). Furthermore, the scale converges as expected with depression scores (r = .78) and other measures of anxiety (r = .73-.85; Spielberger & Reheiser, 2009). We also measured state anxiety using this inventory before and after the trustworthiness appraisal task, which is described in greater detail in Chapter 3.

Because the tasks included here are necessarily social in nature, social anxiety was also evaluated using a self-report form of the Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987). This is a 24-item scale covering fear and avoidance of social situations, including eating in public places, meeting strangers, and telephoning in public. The scale was originally designed to be administered by clinicians, but has satisfactory psychometric properties when administered
via self-report. Internal consistency ($\alpha = .95$) and test-retest reliability (.83 over 12 weeks; Baker, Heinrichs, Kim, & Hofmann, 2002) are both high.

**General Trust and Trustworthiness**

The main measure of trust in the current dissertation was behavioral—participants were asked to view a series of faces and to rate the trustworthiness of each face while also viewing a series of images with neutral, negative, or positive affective valence. Results concerning this measure are presented in Chapter 3. In addition, self-report of general trust, as well as how trustworthy a person considers him/herself to be, were measured. Psychodynamic perspectives suggest that individuals with BPD may have poor self-other differentiation (e.g., Fonagy et al., 2011; Beeney et al., 2015). In other words, individuals who see themselves as less trustworthy may also view others similarly. A short self-report scale querying for general trust was developed and included (see Appendix A). The mistrust subscale of the SNAP-2 (described above) also provided a way to capture general trust.

**Childhood Trauma**

Early experiences of abuse and neglect from family members are common in BPD. Children who experience abuse and neglect are nearly 8 times more likely to develop BPD (Johnson et al., 1999), and up to 76 percent of inpatients with BPD report a history of early sexual or physical abuse (Zanarini, 2000). Childhood abuse is related to aggression, negativism, and irritability in BPD (Morandotti et al., 2013). Participants completed a retrospective self-report of childhood trauma, the Childhood Trauma Questionnaire (CTQ; Berenstein & Fink, 1998), a 28-item measure with 5 subscales: physical abuse, sexual abuse, emotional abuse, physical neglect, and emotional neglect.
**Rejection Sensitivity**

Rejection sensitivity has been shown to mediate the relation between BPD features and trust ratings (Miano et al., 2013). Rejection sensitivity was evaluated using the Adult Rejection Sensitivity Questionnaire (ARSQ; Downey & Feldman, 1996). The ARSQ provides the participant with 9 hypothetical situations and assesses the degree to which the participant would feel concerned about rejection as well as to what degree the individual would expect acceptance versus rejection in each situation. The final scale score is calculated by multiplying the concern/anxiety subscale and the expectation subscale to produce a weighted total. The scale has satisfactory internal consistency ($\alpha = .83$) and test-retest reliability over 3-week ($\alpha = .83$) and 4-month ($\alpha = .73$) intervals (Downey & Feldman, 1996). A typical mean score for student samples ($M = 8.75$, $SD = 3.39$; Taylor & Reeves, 2007) has been reported. The ARSQ predicts relationship dissolution (Downey, Freitas, Michaelis, & Khouri, 1998) and converges with measures of social distress, low self-esteem (Downey & Feldman, 1996), and borderline features (Ayduk et al., 2008).

Participants also completed a second measure of rejection sensitivity: the McLean Assessment of Rejection Sensitivity (MARS; Choi-Kain, unpublished). This scale includes 10 self-report items that assess how readily an individual perceives rejection and rejection vigilance (attention allocation to signs of rejection, rejection avoidance). Items are assessed on a 1-4 scale. Reverse-coded items are both reversed and multiplied by -1, and total scores range from -10 to 20. The scale has strong internal consistency (Masland, unpublished data). Because this measure is unpublished, it is included in Appendix B.
Emotional Reactivity

Emotional reactivity was assessed using the Emotional Reactivity Scale (ERS; Nock, Wedig, Holmberg, & Hooley, 2008), a 21-item self-report measure capturing emotion sensitivity, intensity, and persistence. The scale has good internal consistency (α = .94), as well as divergent, convergent, and criterion validity.

Negative Self-Concept

BPD is associated with a profoundly negative self-concept (Zeigler-Hill & Abraham, 2006). This may be expected to influence appraisals of others, given known difficulty with self-other differentiation (e.g., Fonagy et al., 2011). To explore this idea, participants completed measure of self-criticism and negative self-worth, the Self Rating Scale (SRS; Hooley et al., 2010). The SRS has 8 items and strong internal consistency (α = .73).

Mood

Mood before and after the main behavioral measure of trust processing was assessed using the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). The PANAS is a 20-item scale measuring current experience of mood along positive (e.g., interested, enthusiastic, strong) and negative (e.g., hostile, irritable, scared) dimensions. The scale has high internal consistency (Molloy, Pallant, & Kantas, 2001; Watson, Clark, & Tellegen, 1988), as well as test-retest reliability over a two-month period, good convergent and divergent validity, and uncorrelated factors (Watson et al., 1988).

Social Functioning

Social functioning was measured using three already-validated scales as well as a novel measure. Participants completed the Social Function Questionnaire (SFQ; Tyrer et al., 2005), a brief, 8-item self-report scale developed from a semi-structured interview. The scale has strong
test-retest and inter-rater reliability. However, item-level inspection suggests that some of the specific questions are closely related to symptoms of BPD. In a BPD sample they may tap more into borderline pathology itself rather than general social functioning; this was one consideration in the decision to include additional measures of social functioning. Participants also completed a modified version of the Experiences in Close Relationships Scale-Revised (ECR; Fraley, Waller, & Brennan, 2000). This scale was modified in two ways. First, items were re-worded to refer to relationships in general rather than to romantic relationships specifically. Second, the scale was augmented with additional items intended to assess social functioning. Finally, we included the Social Adjustment Scale-Self Report (SAS; Weissman, 1999). The SAS includes 54 self-report items and assesses an individual’s ability to function and derive satisfaction from various life roles. The assessment of social functioning, including additional psychometric information about the SAS is described in more detail in Chapter 5.

Finally, in addition to the SFQ, ECR, and SAS, participants also completed 3 novel measures of social functioning designed to more fully capture the manner in which social impairment might manifest for individuals with BPD. These measures were pilot tested in Chapter 5 and included in Chapter 6 alongside the validated measures described here to examine how trustworthiness appraisal processes relate to social functioning.

**Sample Characteristics**

Participants were 79 adults from the community. One participant was excluded from analyses due to a pattern of aberrant responding: he selected one extreme item for every question within each scale. This same participant also reported a schizophrenia diagnosis and a history of traumatic brain injury during the lab visit, neither of which were disclosed during phone screening. Another participant was excluded from analyses due to estimated IQ below 80.
(WTAR estimated IQ = 77). The final sample thus included 77 adults ages 18-70 with a mean age of 31.53 years (SD = 14.01). Participants were predominantly female (72.7%) and Caucasian (64.9%). Other ethnicities represented include Black/African American (15.6%), Asian (3.9%), Hispanic (10.4%) and other (5.2%). The majority of the sample had some college education (91.2%), and a minority of participants endorsed a bachelor’s degree (20.8%) or an advanced degree (19.5%). Average years of education was 14.78 (SD = 1.55). The majority of participants were single (71.4%) and heterosexual (74%).

**Group Differentiation and Personality Pathology**

Each participant was placed into one of two groups on the basis of borderline trait endorsement on the SNAP-2. Individuals endorsing 3 or more borderline traits were placed in the borderline group (BG; n = 30), and participants who endorsed 0-2 traits were placed in the control group (CG; n = 47). On the SNAP-2, the borderline group endorsed an average of 4.43 diagnostic criteria for borderline personality disorder (SD = 1.19). The control group endorsed an average of 0.47 diagnostic criteria (SD = 0.78). As expected given the manner in which groups were defined, the borderline group endorsed significantly more borderline symptoms on SNAP-2, \( t(44.69) = 16.14, p < .001 \).

On the SCID-II, the SNAP-derived borderline group endorsed an average of 4.33 diagnostic criteria for borderline personality disorder (SD = 2.43). The control group endorsed an average of 0.87 diagnostic criteria (SD = 1.65). Individuals in the borderline group endorsed significantly more borderline criteria than the control group on the SCID-II, corrected for unequal group variance, \( t(46.12) = 6.87, p < .001 \). Thus the groups were defined by borderline pathology on a self-report measure (SNAP-2), but were also clearly differentiated by borderline pathology as assessed via structured clinical interview.
Rates of endorsement for specific diagnostic criteria are included in Table 2.1, along with information on group differences. The SNAP-derived borderline group endorsed higher rates of each of 9 borderline diagnostic criteria on the SCID-II, relative to the control group. The most commonly endorsed symptom in the borderline group was self-harm (non-suicidal self-injury). This is not surprising given recruitment strategies: self-harm is a behavior commonly seen in treatment settings, and a portion of the participants in the borderline group was referred by community clinicians. Additionally, a portion of participants was referred from another ongoing study in the laboratory that examined self-harm specifically.

Table 2.1. Symptoms of borderline personality disorder in the borderline and control groups.

<table>
<thead>
<tr>
<th>BPD Symptom</th>
<th>SNAP-2</th>
<th>SCID-II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Borderline</td>
<td>Control</td>
</tr>
<tr>
<td></td>
<td>($n=30$)</td>
<td>($n=47$)</td>
</tr>
<tr>
<td>Total Symptoms Endorsed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>3.7</td>
<td>0.2</td>
</tr>
<tr>
<td>Mean</td>
<td>1.19</td>
<td>0.78</td>
</tr>
<tr>
<td>Total Symptoms Endorsed</td>
<td>$\chi^2$</td>
<td>$p$</td>
</tr>
<tr>
<td>Abandonment</td>
<td>14 (46.7)</td>
<td>4 (8.5)</td>
</tr>
<tr>
<td>Unstable Interpersonal Relations</td>
<td>10 (33.3)</td>
<td>5 (10.6)</td>
</tr>
<tr>
<td>Identity Disturbance</td>
<td>11 (36.7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Impulsivity</td>
<td>17 (56.7)</td>
<td>7 (14.9)</td>
</tr>
<tr>
<td>Self-Harm</td>
<td>23 (76.7)</td>
<td>6 (12.8)</td>
</tr>
<tr>
<td>Affective Instability</td>
<td>14 (46.7)</td>
<td>8 (17)</td>
</tr>
<tr>
<td>Emptiness</td>
<td>15 (50)</td>
<td>3 (6.4)</td>
</tr>
<tr>
<td>Anger</td>
<td>10 (33.3)</td>
<td>5 (10.6)</td>
</tr>
<tr>
<td>Dissociation/Paranoia</td>
<td>16 (53.3)</td>
<td>3 (6.4)</td>
</tr>
</tbody>
</table>

Note. Numbers in parantheses are percentages.
As detailed in Table 2.2, group characteristics were also examined on dimensional measures of personality pathology. As expected, the borderline features group scored significantly higher on the dimensional SNAP-2 borderline scale than the control group. This scale is a measure of dimensional differences in borderline pathology rather than a count of criteria endorsed. Groups also differed on antisocial personality, paranoid personality, schizotypal personality, aggression, impulsivity, and self-harm. In all cases the borderline group endorsed greater psychopathology. Groups did not differ on narcissistic personality traits endorsed.

SNAP-derived groups also differed on the PAI Borderline scale and across all subscales (affective instability, identity problems, negative relationships, and self-harm). Additionally, the mean PAI-BOR T score for the borderline features group observed here (70.43) is roughly equivalent to T score cutoffs for this scale that have been shown to capture individuals who meet criteria for BPD. A T score greater than or equal to 70 has shown 73-82% accuracy in capturing borderline diagnoses (Bell-Pringle, Pate, & Brown, 1997; Stein, Pinkser-Aspen, & Hilsenroth, 2007). A T score of 70 is considered a “moderate elevation” on the PAI-BOR (Morey, 1991).

As outlined in Table 2.2, all differences were in the expected direction: the borderline group endorsed higher rates of psychopathology. Notably, across both SNAP and PAI scales, the control group consistently differed from mean standardization samples by less than 1 standard deviation.
Table 2.2. Measures of personality pathology in the borderline and control groups.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Borderline</th>
<th>Control</th>
<th>t</th>
<th>p</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP Borderline PD</td>
<td>72.86</td>
<td>47.52</td>
<td>13.11</td>
<td>&lt;.001</td>
<td>0.696</td>
</tr>
<tr>
<td>SNAP Antisocial PD</td>
<td>56.68</td>
<td>48.04</td>
<td>3.29*</td>
<td>0.002</td>
<td>0.144</td>
</tr>
<tr>
<td>SNAP Narcissistic PD</td>
<td>49.20</td>
<td>47.76</td>
<td>0.58</td>
<td>0.565</td>
<td>0.004</td>
</tr>
<tr>
<td>SNAP Paranoid PD</td>
<td>67.08</td>
<td>47.52</td>
<td>6.26</td>
<td>&lt;.001</td>
<td>0.343</td>
</tr>
<tr>
<td>SNAP Schizotypal PD</td>
<td>64.75</td>
<td>49.27</td>
<td>7.45</td>
<td>&lt;.001</td>
<td>0.425</td>
</tr>
<tr>
<td>SNAP Aggression</td>
<td>59.32</td>
<td>49.68</td>
<td>3.16*</td>
<td>0.003</td>
<td>0.145</td>
</tr>
<tr>
<td>SNAP Impulsivity</td>
<td>53.84</td>
<td>45.07</td>
<td>4.04*</td>
<td>&lt;.001</td>
<td>0.205</td>
</tr>
<tr>
<td>SNAP Self Harm</td>
<td>79.61</td>
<td>48.59</td>
<td>9.90*</td>
<td>&lt;.001</td>
<td>0.622</td>
</tr>
<tr>
<td>PAI-Borderline (BOR)</td>
<td>70.43</td>
<td>52.96</td>
<td>6.45</td>
<td>&lt;.001</td>
<td>0.357</td>
</tr>
<tr>
<td>BOR-A</td>
<td>70.33</td>
<td>52.47</td>
<td>6.62</td>
<td>&lt;.001</td>
<td>0.369</td>
</tr>
<tr>
<td>BOR-I</td>
<td>65.00</td>
<td>52.47</td>
<td>5.17</td>
<td>&lt;.001</td>
<td>0.262</td>
</tr>
<tr>
<td>BOR-N</td>
<td>66.70</td>
<td>53.09</td>
<td>5.02</td>
<td>&lt;.001</td>
<td>0.251</td>
</tr>
<tr>
<td>BOR-S</td>
<td>62.93</td>
<td>51.06</td>
<td>3.45</td>
<td></td>
<td>0.137</td>
</tr>
</tbody>
</table>

*Note.* SNAP and PAI analyses are based on T scores, as these are more directly interpretable than raw scores.

*Levene's test for equality of variances was significant. Corrected statistics are reported.

**Demographic Differences by Group**

As detailed in Table 2.3., individuals in the borderline group were significantly younger than those in the control group, \( t(65.97) = 5.54, p < .01 \) (corrected for unequal group variances).

Groups did not differ on education, \( t(75) = .959, p = .34 \), gender, \( \chi^2(1) = 1.77, p = .138 \), Cramer’s
\( V = .152; \) ethnicity, \( \chi^2(3) = 6.78, p = .079, \) Cramer's \( V = .297; \) or marital status, \( \chi^2(2) = 1.99, p = .370, \) Cramer's \( V = .161. \)

Table 2.3. Demographic characteristics for borderline and control groups.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Borderline M</th>
<th>SD</th>
<th>Control M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>23.2</td>
<td>6.13</td>
<td>36.85</td>
<td>15.05</td>
</tr>
<tr>
<td>Sex (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>80</td>
<td></td>
<td>68.1</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20</td>
<td></td>
<td>31.9</td>
<td></td>
</tr>
<tr>
<td>Education (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar School</td>
<td>3.3</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0</td>
<td></td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>Some College</td>
<td>66.7</td>
<td></td>
<td>46.8</td>
<td></td>
</tr>
<tr>
<td>College Graduate</td>
<td>13.3</td>
<td></td>
<td>25.5</td>
<td></td>
</tr>
<tr>
<td>Postgraduate</td>
<td>16.7</td>
<td></td>
<td>21.3</td>
<td></td>
</tr>
<tr>
<td>Marital Status (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married/Cohabitating</td>
<td>20</td>
<td></td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Separated/Divorced</td>
<td>3.3</td>
<td></td>
<td>12.8</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>76.7</td>
<td></td>
<td>70.2</td>
<td></td>
</tr>
<tr>
<td>Ethnicity (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>80</td>
<td></td>
<td>55.3</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>3.3</td>
<td></td>
<td>23.4</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>10</td>
<td></td>
<td>10.6</td>
<td></td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>6.7</td>
<td></td>
<td>10.6</td>
<td></td>
</tr>
</tbody>
</table>
Intelligence Estimates

Intelligence was estimated in both groups using the WTAR (described above). Average estimated IQ for the sample overall was 108.30 (SD = 10.85). As shown in Table 2.4, the borderline group had a significantly higher IQ than the control group.

Table 2.4. WTAR intelligence estimates by group

<table>
<thead>
<tr>
<th></th>
<th>Borderline</th>
<th>Control</th>
<th>t*</th>
<th>p</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTAR Estimated IQ</td>
<td>112.07</td>
<td>105.89</td>
<td>2.79</td>
<td>0.007</td>
<td>0.085</td>
</tr>
</tbody>
</table>

*Levene's test revealed unequal variances across groups. Corrected statistic reported.

Other Clinical/Psychological Characteristics

Relative to the control group, the borderline group showed elevated levels of psychopathology and maladaptive psychological/personality traits. As detailed in Table 2.5, tests of group differences indicated that the borderline group endorsed higher levels of depression, trait and social anxiety, schizotypy, self-criticism/negative self-concept, rejection sensitivity, emotional reactivity, negative temperament, mistrust, and disinhibition (all ps < .01). The borderline group endorsed lower levels of positive temperament relative to the control group. These results are largely consistent with the high levels of negative affect and psychopathology that commonly co-occur with borderline personality.
Table 2.5. Clinical characteristics for borderline and control groups.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Borderline</th>
<th>Control</th>
<th>t</th>
<th>p</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Depression (BDI)</td>
<td>21.07</td>
<td>15.06</td>
<td>5.91</td>
<td>7.00</td>
<td>5.17* &lt;.001</td>
</tr>
<tr>
<td>State Anxiety (STAI)</td>
<td>55.97</td>
<td>11.68</td>
<td>38.02</td>
<td>11.54</td>
<td>6.62 &lt;.001</td>
</tr>
<tr>
<td>Social Anxiety (LSAS)</td>
<td>108.70</td>
<td>29.81</td>
<td>86.02</td>
<td>25.30</td>
<td>3.58  0.001</td>
</tr>
<tr>
<td>Rejection Sensitivity (ARSQ)</td>
<td>11.29</td>
<td>4.90</td>
<td>7.84</td>
<td>4.37</td>
<td>3.22  0.002</td>
</tr>
<tr>
<td>Rejection Sensitivity (RS)</td>
<td>7.70</td>
<td>5.45</td>
<td>2.49</td>
<td>6.07</td>
<td>3.82  &lt;.001</td>
</tr>
<tr>
<td>Schizotypy (SPQ)</td>
<td>27.33</td>
<td>14.68</td>
<td>15.02</td>
<td>14.03</td>
<td>3.69  &lt;.001</td>
</tr>
<tr>
<td>SPQ Referential Thinking</td>
<td>2.27</td>
<td>2.12</td>
<td>1.43</td>
<td>1.99</td>
<td>1.77  0.081</td>
</tr>
<tr>
<td>SPQ Social Anxiety</td>
<td>5.13</td>
<td>2.57</td>
<td>2.81</td>
<td>2.56</td>
<td>3.88  &lt;.001</td>
</tr>
<tr>
<td>SPQ Magical Thinking</td>
<td>1.13</td>
<td>1.57</td>
<td>0.72</td>
<td>1.19</td>
<td>1.30  0.198</td>
</tr>
<tr>
<td>SPQ Unusual Perceptions</td>
<td>1.80</td>
<td>1.86</td>
<td>0.81</td>
<td>1.10</td>
<td>2.64* 0.012</td>
</tr>
<tr>
<td>SPQ Odd Behavior</td>
<td>2.73</td>
<td>2.35</td>
<td>1.32</td>
<td>1.84</td>
<td>2.80* 0.007</td>
</tr>
<tr>
<td>SPQ No Close Friends</td>
<td>3.77</td>
<td>2.71</td>
<td>2.21</td>
<td>2.22</td>
<td>2.75  0.008</td>
</tr>
<tr>
<td>SPQ Odd Speech</td>
<td>4.37</td>
<td>2.57</td>
<td>2.00</td>
<td>2.45</td>
<td>4.06  &lt;.001</td>
</tr>
<tr>
<td>SPQ Constricted Affect</td>
<td>3.03</td>
<td>2.30</td>
<td>1.81</td>
<td>2.06</td>
<td>2.43  0.017</td>
</tr>
<tr>
<td>SPQ Suspiciousness</td>
<td>3.10</td>
<td>2.20</td>
<td>1.91</td>
<td>2.24</td>
<td>2.28  0.026</td>
</tr>
<tr>
<td>Loneliness</td>
<td>49.27</td>
<td>12.80</td>
<td>40.00</td>
<td>13.55</td>
<td>2.99  0.004</td>
</tr>
<tr>
<td>Self-Criticism (SRS)</td>
<td>33.03</td>
<td>13.59</td>
<td>19.49</td>
<td>10.25</td>
<td>4.97  &lt;.001</td>
</tr>
<tr>
<td>Emotional Reactivity (ERS)</td>
<td>46.87</td>
<td>23.04</td>
<td>24.51</td>
<td>19.85</td>
<td>4.53  &lt;.001</td>
</tr>
<tr>
<td>Positive Temperament (SNAP)</td>
<td>14.63</td>
<td>6.37</td>
<td>18.91</td>
<td>6.24</td>
<td>2.91  0.005</td>
</tr>
<tr>
<td>Negative Temperament (SNAP)</td>
<td>21.87</td>
<td>4.66</td>
<td>7.96</td>
<td>6.57</td>
<td>10.85* &lt;.001</td>
</tr>
<tr>
<td>Mistrust (SNAP)</td>
<td>11.37</td>
<td>4.75</td>
<td>5.11</td>
<td>4.20</td>
<td>6.06  &lt;.001</td>
</tr>
<tr>
<td>Disinhibition (SNAP)</td>
<td>13.47</td>
<td>6.01</td>
<td>6.85</td>
<td>4.55</td>
<td>5.48  &lt;.001</td>
</tr>
</tbody>
</table>

*Levene's test for equality of variances was significant. Corrected statistics are reported.
Diagnostic Characteristics

Consistent with elevations across personality measures, as well as high rates of diagnostic comorbidity observed in borderline personality disorder, individuals in the borderline group endorsed higher rates of psychiatric diagnoses (diagnostic frequencies by group are provided in Table 2.6). Although cell sizes for each diagnosis are not large enough across groups to permit fully interpretable tests of significant differences, it is readily apparent that individuals in the borderline group endorsed higher rates of depression (both major and persistent), agoraphobia, social phobia, drug/alcohol abuse/dependence, bulimia, and generalized anxiety disorder. The most commonly endorsed diagnoses within this group included major depression and persistent depressive disorder. This is congruent with past work showing that depression is highly prevalent in BPD (see Köhling et al., 2015). The control group, by contrast, met criteria for relatively few psychiatric diagnoses. This is likely in part attributable to the recruitment strategy employed, in which participants “with no history of mental illness” were sought.
<table>
<thead>
<tr>
<th>Disorder</th>
<th>Borderline</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n =30 )</td>
<td>( n =47 )</td>
</tr>
<tr>
<td>Major Depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>9 (30)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Recurrent</td>
<td>7 (23.3)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Persistent Depressive Disorder</td>
<td>8 (26.7)</td>
<td>2 (4.3)</td>
</tr>
<tr>
<td>Manic Episode</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Past</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Hypomaniac Episode</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Past</td>
<td>5 (16.7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>1 (3.3)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>8 (26.7)</td>
<td>2 (4.3)</td>
</tr>
<tr>
<td>Social Phobia</td>
<td>7 (23.3)</td>
<td>2 (4.3)</td>
</tr>
<tr>
<td>Obsessive-Compulsive Disorder</td>
<td>1 (3.3)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Posttraumatic Stress Disorder</td>
<td>2 (6.7)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Alcohol Dependence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past Year</td>
<td>2 (6.7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Lifetime</td>
<td>1 (3.3)</td>
<td>2 (4.3)</td>
</tr>
<tr>
<td>Alcohol Abuse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past Year</td>
<td>3 (10.0)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Lifetime</td>
<td>1 (3.3)</td>
<td>2 (4.3)</td>
</tr>
<tr>
<td>Drug Abuse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past Year</td>
<td>3 (10.0)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Lifetime</td>
<td>1 (3.3)</td>
<td>1 (2.1)</td>
</tr>
<tr>
<td>Drug Dependence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past Year</td>
<td>4 (13.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Lifetime</td>
<td>1 (3.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Anorexia</td>
<td>2 (6.7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Bulimia</td>
<td>5 (16.7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Generalized Anxiety Disorder</td>
<td>8 (26.7)</td>
<td>1 (2.1)</td>
</tr>
</tbody>
</table>

*Note.* All time periods are current unless otherwise specified.
**Childhood Trauma and Parenting**

Individuals with borderline personality disorder or borderline features are known to endorse high rates of childhood trauma (Zanarini, 2000), and the experience of childhood trauma influences how individuals appraise the characteristics of others (e.g., Nicol et al., 2013), including trustworthiness (e.g., Ebert et al., 2013). Childhood trauma is a putative etiological factor in BPD (Zanarini et al., 1997). Children who experience abuse and neglect are nearly 8 times more likely to develop BPD (Johnson et al., 1999), but the experience of abuse/trauma is not common to all individuals with BPD. Furthermore, subtler forms of childhood maltreatment, including emotional neglect or invalidation, may contribute to the etiology of BPD (Linehan, 1993). As outlined in Table 2.7, individuals in the borderline group endorsed higher rates of emotional abuse, emotional neglect, and physical neglect (trending toward significant). However, they did not differ from the control group on physical abuse or sexual abuse. Differences between groups are potentially magnified given the finding that more individuals in the control group seemed to minimize childhood trauma more frequently than the borderline group, as measured by the minimizing/denying subscale of the CTQ. However, it is unclear whether this might be due to a magnification effect in the borderline features group or a minimization effect in the control group.

Similarly, individuals in the borderline group endorsed higher rates of problematic parenting from their mothers and fathers than the control group on the MOPS. The MOPS captures aspects of problematic parenting that may not rise to the full level of neglect or abuse. Additionally, it includes two abuse subscales. Further information, including group means, is presented in Table 2.7.
Here I have described recruitment and experimental procedures. Participants were recruited from the community, from outpatient clinical settings, and from other studies ongoing at Harvard University. Recruitment used a “take all” approach with few exclusionary criteria and attempts to target individuals with borderline personality features. Participants were divided into two groups (borderline vs. control) on the basis of trait endorsement on the SNAP-2. This strategy was successful. SNAP-derived groups differed significantly and in the expected direction on multiple measures of borderline pathology, including the SCID-II and the PAI. They also differed, as expected, across other measures of psychopathology: the borderline group consistently endorsed more severe and more frequent psychopathology, including depression and

<table>
<thead>
<tr>
<th>Measure</th>
<th>Borderline</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Childhood Trauma Questionnaire</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional Abuse</td>
<td>14.67</td>
<td>6.23</td>
</tr>
<tr>
<td>Physical Abuse</td>
<td>8.47</td>
<td>4.51</td>
</tr>
<tr>
<td>Sexual Abuse</td>
<td>5.20</td>
<td>2.66</td>
</tr>
<tr>
<td>Physical Neglect</td>
<td>9.23</td>
<td>3.91</td>
</tr>
<tr>
<td>Emotional Neglect</td>
<td>15.13</td>
<td>5.90</td>
</tr>
<tr>
<td>Minimizing/Denying</td>
<td>2 (6.7)</td>
<td>19 (36.2)</td>
</tr>
<tr>
<td>Measure of Parenting Style</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother Indifference</td>
<td>5.87</td>
<td>5.45</td>
</tr>
<tr>
<td>Mother Abuse</td>
<td>4.93</td>
<td>4.88</td>
</tr>
<tr>
<td>Mother Overcontrol</td>
<td>6.87</td>
<td>3.79</td>
</tr>
<tr>
<td>Father Indifference</td>
<td>7.37</td>
<td>6.90</td>
</tr>
<tr>
<td>Father Abuse</td>
<td>6.23</td>
<td>6.53</td>
</tr>
<tr>
<td>Father Overcontrol</td>
<td>5.77</td>
<td>3.81</td>
</tr>
</tbody>
</table>

Note. Minimizing/denying values are counts and (percentages).
*Levene's test for equality of variances was significant. Corrected statistics are reported.

Summary

Here I have described recruitment and experimental procedures. Participants were recruited from the community, from outpatient clinical settings, and from other studies ongoing at Harvard University. Recruitment used a “take all” approach with few exclusionary criteria and attempts to target individuals with borderline personality features. Participants were divided into two groups (borderline vs. control) on the basis of trait endorsement on the SNAP-2. This strategy was successful. SNAP-derived groups differed significantly and in the expected direction on multiple measures of borderline pathology, including the SCID-II and the PAI. They also differed, as expected, across other measures of psychopathology: the borderline group consistently endorsed more severe and more frequent psychopathology, including depression and
anxiety. This group also endorsed higher scores on measures of processes that putatively support psychopathology (e.g., negative temperament) and more significant history of childhood emotional abuse and experiences of negative parenting. Notably, groups also differed significantly on IQ and age. The borderline group had a significantly higher IQ and was significantly younger than the control group. These results were not expected, and are an important consideration in the evaluation of the study results that follow in the coming chapters.
Chapter 3: Affective Priming and Trustworthiness Appraisal
Introduction

For most individuals, daily life requires interaction with strangers, acquaintances, and close others. The ability to navigate these interactions successfully relies on a host of cognitive, emotional, and behavioral processes. Consider that at a minimum an interaction requires you to evaluate both verbal and non-verbal cues, interpret ambiguities in speech and facial expression, modulate your emotional experience, and respond appropriately. Because social interaction is inherently complex, humans use a wealth of shortcuts to infer the mental states, traits, and intentions of others. Implicit trait appraisal is a common heuristic used to navigate social interactions. Based on very little information, humans quickly make judgments about the traits of other individuals, including how dominant, likeable, intelligent, or trustworthy they are.

These quickly formed judgments allow humans to make predictions about the future actions of others and to infer their mental states (e.g., beliefs, emotions, thoughts, motivations). An initial impression that someone is dominant and untrustworthy, for example, may lead an individual to protectively provide that person with less personal information or fewer resources. Indeed, ratings of trustworthiness predict voting behaviors and election outcomes (Todorov et al., 2005), and negative attributions (e.g., low trustworthiness) about candidates are more important than positive attributions (e.g., attractiveness; Spezio et al., 2008). This is perhaps not surprising when one considers that judging another individual as trustworthy relies in part on imagining their future interpersonal behaviors (Freedman et al., 1951; Gilbert & Malone, 1995).

Prior research indicates that trait appraisal is biased for some groups of individuals, including those with schizophrenia (e.g., Hooker et al., 2011) and borderline personality. Individuals with borderline personality disorder and subclinical borderline traits rate faces as less trustworthy in laboratory paradigms (Nicol et al., 2013; Fertuck et al., 2013; Miano et al., 2013). Evidence suggests that this is due to a response bias, rather than to heightened sensitivity to
objective cues that an individual is untrustworthy or to greater ability to discriminate trustworthy versus untrustworthy individuals (Fertuck et al., 2013). Importantly, Fertuck et al. (2013) also found that this trust bias occurs independently of fear appraisal, indicating that less trustworthy faces are not simply those that evoke greater fear response. Using a similar paradigm to Fertuck et al. (2013) with both facial and limited behavioral information available to subjects, Arntz and Veen (2001) found that individuals with BPD rate others more negatively in general when viewing short movie clips. Relatedly, Barnow et al. (2009) found that such individuals rate others as more aggressive than do either healthy controls or depressed patients. Individuals with BPD also have a bias to rate others as unapproachable, and this bias may be best explained by childhood trauma for individuals with BPD (Nicol et al., 2013).

Despite the knowledge that both interpersonal dysfunction and affective instability are central to BPD, and evidence suggesting that the two interact in important ways, no work to date has examined how affective arousal might influence trust appraisal for individuals with borderline traits. This gap is particularly striking in light of extant knowledge implicating the amygdala, a structure central to affective processes, in the processing of trait appraisals including trustworthiness.

Affective Processing and Trait Appraisals

Trait appraisal is intimately tied to affective processes. Even when participants are not actively tasked with evaluating emotional information, the amygdala responds to novel faces according to general valence (of which trustworthiness is a part), and automatically modulates activation of cortical areas important for face perception (Todorov & Engell, 2008). Thus although the systems serving trait appraisal (e.g., mostly frontal circuits; Bechera et al., 2000) are more flexible and partly dissociable from those that serve emotion recognition (e.g., sensory
cortex; Haxby, Hoffman, & Gobbini, 2002), affective/emotional information directly influences neural processing of trait information. The amygdala automatically processes threatening information quickly, efficiently, and involuntarily (Phelps & LeDoux, 2005; Moors & De Houwer, 2006). Cues of untrustworthiness are inherently a signal of social threat; untrustworthy individuals are more likely to be rejecting, dishonest, negatively judgmental, and emotionally harmful (Fertuck, 2013). Identifying untrustworthy individuals is functionally protective and may guide social behavior. Given evidence that amygdala function has evolved in tandem with the evolution of human social behavior (Amaral et al., 2003), one may rationally expect the amygdala to have a role in evaluating the trustworthiness of others. In fact, greater amygdala activation to negatively valenced or untrustworthy faces may mediate allocation of attentional resources to this information, and may influence memory for faces (Todorov & Engell, 2008).

Unbiased trait appraisal requires individuals to modulate their own emotional responses and control the influence of their affective experience on their appraisals. Individuals who have more difficulty modulating amygdala responses to facial stimuli or cues of trustworthiness may over-perceive threat, resulting in lower ratings of trustworthiness. Biased trustworthiness ratings in BPD might therefore be mediated by dysregulation in frontolimbic networks responsible for modulating response to social threat information.

**Amygdala and BPD**

Dysregulated neural processing in frontolimbic networks in borderline personality disorder may be driven in large part by amygdala dysfunction and may be understood as an essential part of the disorder’s pathophysiology; this dysregulation may be responsible for the impulsivity and affective lability that characterize the disorder (Hoerst et al., 2010; Schmahl & Bremner, 2006). Connectivity of the prefrontal cortex with limbic regions including the
Amygdala reactivity is greater in borderline personality disorder relative to healthy controls. When shown repeated emotional information (but not neutral information), individuals with BPD show greater and more prolonged activation (i.e., less habituation) in the amygdala relative to both healthy controls and individuals with schizotypal personality disorder (Hazlett et al., 2012) or avoidant personality disorder (Koenigsberg et al., 2014). Individuals with BPD also show greater amygdala reactivity to stimuli of various types including emotional pictures (Herpertz et al., 2001), emotional scripts (Beblo et al., 2006), and emotional faces (Donegan et al., 2003). Recent conceptualizations of the pathophysiology underlying affective lability in BPD posit that over-reactive and sensitive amygdala activation with impaired habituation are likely causal. Initial evidence suggests that improved amygdala habituation may be a mechanism for improving emotion regulation following dialectical behavior therapy (Goodman et al., 2014).

Greater amygdala activity across types of stimuli and situations suggests that an essential part of the pathophysiology of borderline personality is emotional over-reactivity. Neural over-reactivity to emotional stimuli may be one mechanism that contributes to the difficulties regulating emotion that are both observable and apparent on self-report. Furthermore, difficulty controlling the influence of emotional stimuli may contribute to biases in social cognition, including how individuals with borderline personality attribute trustworthiness to others. Here we sought to
examine whether affective priming exerted undue influence over trustworthiness appraisals for individuals with borderline personality relative to controls.

**Rejection Sensitivity and Trust Appraisal Bias**

Individuals with borderline personality are hypersensitive to rejection (Staebler, Helbing, Rosenbach, & Renneberg, 2011), and show greater emotional and physiological arousal to stimuli that evoke themes of rejection (Limberg et al., 2011; Schmahl et al., 2004). On an interpersonal level, individuals with borderline personality express an intense need for closeness with others and an equally intense fear of rejection (Gunderson & Ruth-Lyons, 2008). Rejection sensitivity is so intimately tied to our conceptualizations of borderline personality and so apparent in interactions with individuals with the disorder that some have asserted that it may be a core psychological endophenotype for borderline pathology (e.g., Gunderson, 2007).

Rejection sensitivity is conceptualized as a cognitive-affective tendency to perceive rejection easily, overreact to rejection emotionally, and to expect rejection anxiously (Downey & Feldman, 1996). Rejection sensitivity is conceptually similar to fearful attachment, which is in turn related to more negative trait appraisals (Horppu & Ikonen-Varila, 2001). Individuals with borderline personality endorse higher levels of rejection sensitivity than both healthy controls and individuals with social anxiety disorder (Staebler, Helbing, Rosenbach, & Renneberg, 2011). Additionally, individuals high on rejection sensitivity are more likely to make negative attributions about partner behavior (Downey & Feldman, 1996). This indicates that rejection sensitivity has important implications for social cognitive processing.

Ascribing trustworthiness to an individual implies an expectation that the individual will not act in a rejecting manner. Individuals who are more perceptually sensitive or vigilant to rejection may therefore be more likely to view others as less trustworthy. Thus it may be those
individuals with borderline personality and high rejection sensitivity who are most likely to exhibit negative trust appraisal biases. Indeed, prior research has found that rejection sensitivity mediates the association between borderline features and negative trust appraisal bias (Miano et al., 2013). Here we sought to replicate this finding and to determine how rejection sensitivity interacts with affective priming to influence trait appraisals.

**Childhood Trauma and Trust Appraisal Bias**

Many individuals with borderline personality disorder or traits may have good reason to fear rejection. High rejection sensitivity may be the result of childhood patterns of rejection, neglect, and exclusion (Downey & Feldman, 1996). By this conceptualization, rejection sensitivity originates in childhood and is likely an adaptive disposition; attending to negative social cues could protect an individual from emotional investment in an individual who will ultimately be rejecting. Rates of childhood abuse and neglect are high for individuals with borderline traits. Children who experience abuse and neglect are nearly 8 times more likely to develop BPD (Johnson et al., 1999), and up to 76 percent of inpatients with BPD report a history of early sexual or physical abuse (Zanarini, 2000). Abuse from loved ones represents an extreme form of rejection, often from those a child should most be able to trust: close family members.

Other work suggests that the experience of childhood trauma may play a role in trustworthiness appraisals only in interaction with attachment activation. For example, it is now a replicated finding that individuals with BPD act as though others are less trustworthy when administered oxytocin during behavioral trust games (Bartz et al., 2011; Ebert et al., 2013). The trust-lowering effect of oxytocin is augmented for individuals with a history of childhood trauma, while childhood trauma does not predict behavior above and beyond BPD diagnosis when oxytocin has not been administered (Ebert et al., 2013).
Childhood trauma is a type of betrayal trauma in which a close other is a perpetrator of abuse or neglect. Betrayal trauma is associated with lower general trust in relationships, although the experience of betrayal trauma itself may not predict behavior in economic exchange games (Gobin & Freyd, 2014). Individuals with early betrayal trauma generally show high levels of distrust (DiLillo & Long, 1999; Jurgens, 2005; Lau & Kristensen, 2010). Women with a history of childhood abuse have lower concentrations of oxytocin, an affiliative neuropeptide, in their cerebrospinal fluid (Heim et al., 2009). However, alterations in trustworthiness judgments are not always negative for individuals with betrayal trauma; some may exhibit a positive trustworthiness bias in which they are too eager to trust others or may have difficulty identifying dishonesty, while some may exhibit a negative trustworthiness bias (Gobin & Freyd, 2009; Zurbriggen & Freyd, 2004). Thus although it seems clear that betrayal trauma influences trustworthiness appraisals, it is currently unclear what predicts the direction of its influence. It may be the case that some individuals display a bias to attend to happiness cues that overrides the effects of borderline personality or betrayal trauma (Fani et al., 2012). This attentional strategy may only be deployed by a subset of individuals who have experienced betrayal trauma and may help account for the differences in trustworthiness appraisals that result.

Here I predicted that the experience of childhood trauma would moderate the relationship between borderline features and trustworthiness appraisals.
Summary and Predictions

In summary, borderline personality disorder has been associated with biased trustworthiness appraisal. I expected to replicate this finding and predicted that relative to the control group, the borderline group would make significantly lower appraisals of trustworthiness regardless of condition during an affective priming task. Additionally, borderline personality disorder is characterized by significant affective lability and deficits in frontolimbic modulation of emotional responding. Affective lability and oversensitive and under-regulated amygdala activation have been shown to interfere with a variety of cognitive processes for individuals with borderline pathology. These affective processing impairments are likely to interfere similarly with social cognitive processes, including trustworthiness appraisal. More specifically, affective information is likely to exert exaggerated influence on social cognitive judgments for individuals with borderline personality features relative to individuals without borderline pathology. Accordingly, I predicted that (1) relative to the control group, the influence of negative affective primes on trustworthiness appraisals would be significantly greater for individuals in the borderline group; and (2) affective instability would predict priming effects. Finally, borderline personality disorder is often associated with high rejection sensitivity and the experience of childhood trauma. Both rejection sensitivity and the experience of childhood trauma may influence trustworthiness perceptions. I predicted that (1) rejection sensitivity would mediate the relationship between borderline features and priming effects; and (2) the experience of childhood trauma would moderate the relationship between borderline features and trustworthiness appraisals. To test these hypotheses, the current study included an experimental paradigm that required participants to rate the trustworthiness of faces after a series of affective primes.
**Method**

Participants completed a variant of the trustworthiness task designed by Hooker et al. (2011). This computerized social judgment task was adapted from a task developed by Adolphs, Tranel, and Damasio (1998). Participants viewed a series of scenes, each of which was followed by the presentation of a face. Emphasis was placed on the idea that the scenes and faces were not related. Participants were asked to rate the trustworthiness of the faces on a 7-point scale, with -3 corresponding to “very untrustworthy” and 3 corresponding to “very trustworthy.” Following Hooker et al. (2011) and Adolphs et al. (1998), participants were asked to “imagine trusting the person in a very serious situation, for instance, with all your money or with your life.” In order to capture a fuller range of situations in which trust might be important, additional text was added to this statement: “or with an important secret.”

The faces used in the task included 50 black-and-white photographs of unfamiliar male and female faces from the 100-face stimulus set used by Adolphs et al. (1998). Facial expressions were natural and neutral. Normed trustworthiness ratings ranged from -2.45 to 1.57. The faces were preceded by the presentation of affective primes (negative/threatening, neutral, or positive images) taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005). Fifty pictures were selected for each condition. These were the same photographs used by Hooker et al. (2011), who had undergraduates identify the most threatening but least disgusting pictures for the negative condition, and who determined that the neutral primes were sufficiently neutral in valence and low in arousal; that the positive primes were positive in valence and high on arousal significantly more so than the neutral primes; and that the negative primes were significantly more unpleasant and arousing than neutral primes. Neutral primes included mostly pictures of household objects; negative primes involved images of snakes, spiders, weapons, and assault; and positive primes featured sports and food. Primes were
randomly assigned to faces, which were then presented in a pseudo-random fixed order: no faces appeared twice in a row. Primes were presented for 1s, followed by presentation of a face for 7s.

**Priming Effect Index**

Priming effect index was calculated using negative difference scores. Negative differences scores were calculated by subtracting trustworthiness ratings after the neutral prime from trustworthiness ratings after the negative affective prime. Similarly, trustworthiness ratings after the neutral prime were subtracted from ratings after the positive prime to calculate positive difference scores. In this way I controlled for within-subjects baseline trustworthiness ratings in order to index the effect of affective priming.

**Other Measures**

Measures are described in detail in Chapter 2. Relevant to the examination of affective priming influences on trustworthiness appraisals, I measured rejection sensitivity using the Adult Rejection Sensitivity Questionnaire (ARSQ; Berenson et al., 2009) and the McLean Assessment of Rejection Sensitivity (MARS; Choi-Kain, unpublished). Childhood trauma exposure was assessed using the Childhood Trauma Questionnaire (CTQ; Berenstein & Fink, 1997). I also measured positive and negative affect before and after the trust task using the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), as well as trait and state anxiety before and after the task using the State-Trait Anxiety Inventory (STAI; Spielberger, 1983).
Results

Affective Prime Validation

If the affective primes effectively influenced trustworthiness appraisals, both groups should demonstrate lowest trustworthiness ratings in the negative priming condition and highest trustworthiness ratings in the positive priming condition. Results demonstrated that both groups were significantly influenced by affective priming. This indicates that the priming paradigm was effective. The borderline group rated faces as significantly less trustworthy after the negative prime ($M = -0.47, SD = 0.63$) relative to the neutral prime ($M = -0.21, SD = 0.57; t(29) = 3.15, p = .004$). A similar trend was observed in the control group, who rated faces as less trustworthy following negative primes ($M = 0.09, SD = 0.74$) as compared to neutral primes ($M = 0.16, SD = 0.75; t(46) = 1.84, p = .072$). The control group also rated faces as significantly more trustworthy following positive primes ($M = 0.26, SD = 0.75$) as compared to neutral primes ($t(46) = 3.91, p < .001$). A similar significant effect was observed in the borderline group (following positive primes $M = -0.63, SD = 0.59; t(29) = 3.38, p = .002$).

Negative and Positive Mood and State Anxiety

Recent evidence suggests that trait anxiety is a significant and specific predictor of trustworthiness ratings and must be accounted for as research seeks to build models of the neural and cognitive mechanisms underlying this type of appraisal (Willis, Dodd, & Palermo, 2013). Here I evaluated participants on trait anxiety using the State-Trait Anxiety Inventory (STAI; Spielberger, 1983). Chapter 2 provides a full description of this measure. I also used this measure to examine state anxiety, with the prediction that both trait and state anxiety would make significant contributions to trustworthiness appraisals. Because social anxiety can play a modulating role in how individuals react to subliminal face primes (Paul, Pope, Fennell, & Mendl, 2012), we measured this construct using the Liebowitz Social Anxiety Scale (LSAS;
Liebowitz, 1987), also described in Chapter 2. Borderline features were significantly associated with both trait and social anxiety. Intercorrelations among state/trait anxiety, mood, and borderline features measures are presented in Table 3.1. Group differences on trait and social anxiety are presented in Table 3.2. Although the borderline and control groups differ on these and other measures of psychopathology, it is not appropriate to use ANCOVA to statistically “control” for the effect of psychopathology other than borderline features in the group analyses that follow. Statistical control for covariates between non-randomized groups is not appropriate (Miller & Chapman, 2001). Because borderline pathology is in part characterized by high levels of negative affect, removing variance associated with anxiety or depression would remove variance that is inextricably tied to the defining features of the groups.

Measures of state and trait anxiety and mood were largely unrelated to raw trustworthiness appraisals and priming effects. Correlations of these measures with trustworthiness scores are summarized in Table 3.3. Two exceptions emerged: positive difference scores (a measure of the influence of positive primes) were significantly related to state anxiety change and negative mood change, such that the effect of positive primes was diminished as anxiety and negative mood increased.

Furthermore, participants completed the state anxiety items of the STAI before and after the main trust appraisal, as well as the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). These measures were included before and after the task to rule out the possibility that changes in mood or state anxiety could account for any observed differences in trustworthiness appraisals or priming effects. Other work has shown that state negative affect is not associated with trust bias (Fertuck et al., 2013), but whether state negative affect influences the effect of priming on trust bias has not been explored.
Intercorrelations among borderline pathology, negative and positive mood change, state anxiety change, trait anxiety, and social anxiety are displayed in Table 3.1. Borderline features (measured both by number of disorder criteria endorsed and a normed T score on the SNAP) were not related to any state changes in anxiety, positive mood, or negative mood. Furthermore, the borderline and control groups did not differ on any of these variables when tested categorically with independent samples t tests. Results of group comparisons are included in Table 3.2. Notably, this is not due to a ceiling effect. Although the borderline group endorsed a higher level of state anxiety at baseline \((M = 40.03, SD = 11.57)\) than the control group \((M = 28.53, SD = 7.44; t(75) = 5.32, p < .001)\), this level was not near the highest possible score (80) on this measure. State anxiety scores for the borderline group after the task \((M = 42.20, SD = 10.96)\) were also not approaching the measure’s ceiling. We can thus have reasonable confidence moving forward that any effects found for the borderline group are not merely attributable to changes in mood or anxiety.

<p>| Table 3.1. Intercorrelations of state and trait anxiety/mood measures and borderline features. |
|-------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>SNAP Borderline Criteria</th>
<th>SCID-II Borderline Criteria</th>
<th>SNAP Borderline T Score</th>
<th>Trait Anxiety</th>
<th>State Anxiety Change</th>
<th>Social Anxiety</th>
<th>Positive Mood Change</th>
<th>Negative Mood Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP Borderline Criteria</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>SCID-II Borderline Criteria</td>
<td><strong>.723</strong></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>SNAP Borderline T Score</td>
<td><strong>.924</strong></td>
<td><strong>.763</strong></td>
<td><strong>.654</strong></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Trait Anxiety</td>
<td><strong>.701</strong></td>
<td><strong>.619</strong></td>
<td><strong>.654</strong></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>State Anxiety Change</td>
<td>-.021</td>
<td>-.037</td>
<td>-.012</td>
<td>-.094</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Social Anxiety</td>
<td><strong>.442</strong></td>
<td><strong>.383</strong></td>
<td><strong>.478</strong></td>
<td><strong>.623</strong></td>
<td><strong>.075</strong></td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Positive Mood Change</td>
<td>-.084</td>
<td>-.203</td>
<td>-.116</td>
<td>-.072</td>
<td><strong>.363</strong></td>
<td><strong>.013</strong></td>
<td>--</td>
</tr>
<tr>
<td>Negative Mood Change</td>
<td>-.042</td>
<td>-.037</td>
<td>.01</td>
<td>.006</td>
<td><strong>.674</strong></td>
<td>-.046</td>
<td><strong>.341</strong></td>
</tr>
</tbody>
</table>

*Note. Values in table are Pearson correlations, with the exception of all correlations with SNAP Borderline Criteria and SCID Criteria, which are Spearman rho correlations.

**p < .01, *p < .05 (two-tailed tests)
Self-Reported Trust and Trustworthiness

Self-report of general trust was significantly lower in the borderline group ($M = 19.91$, $SD = 5.45$) than in the control group ($M = 22.85$, $SD = 4.64$; $t(75) = 2.48$, $p = .015$). Similarly, report of self-trustworthiness was significantly lower in the borderline group ($M = 28.40$, $SD = 5.98$) than in the control group ($M = 32.02$, $SD = 3.29$; $t(75) = 3.43$, $p = .001$). Individuals with
borderline features both reported that they are less trusting of other people and endorsed being less trustworthy themselves. For the overall sample, self-reported trust was significantly correlated with average trustworthiness ratings such that those who reported trusting other people more also rated faces as more trustworthy, \( r(76) = 0.26, p = .021 \). Self-reported trust was not related to negative \( (r(76) = -0.30, p = .80) \) or positive difference scores \( (r(76) = -0.08, p = .52) \). The degree to which individuals endorsed that they themselves were trustworthy was significantly correlated with average trustworthiness ratings, such that higher self-trustworthiness was associated with more trusting appraisals, \( r(76) = 0.34, p = .002 \). Self-trustworthiness was unrelated to negative \( (r(76) = 0.12, p = .30) \) or positive difference scores \( (r(76) = -0.22, p = .85) \).

Higher self-reported trust was significantly associated with greater belief in one’s own trustworthiness, \( r(76) = 0.31, p = .007 \). Overall, these results suggest that responses on the trustworthiness task accorded with individuals’ own beliefs about how trusting they are. Results also suggest that individuals who believe themselves to be less trusting are also less trusting of others.

**Hypothesis 1**

Relative to the control group, the borderline group will make significantly lower appraisals of trustworthiness regardless of priming condition.

I examined general trust appraisal bias in the borderline personality group with an independent samples t-test. Average trust ratings were calculated for each participant, collapsed across stimulus presentation and all priming conditions. Overall, individuals in the borderline personality group made significantly lower ratings of trustworthiness \( \bar{t}(75) = -2.71, p = .008, R^2 = .089 \). Consistent with past work, this demonstrates that the borderline personality group had a
response bias to rate faces as less trustworthy than the control group, regardless of affective prime condition. Results are included in Table 3.4 and illustrated in Figure 3.1.

Table 3.4. Trust appraisal averages by group and condition.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Borderline</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Neutral Primes</td>
<td>-0.213</td>
<td>0.567</td>
</tr>
<tr>
<td>Negative Primes</td>
<td>-0.472</td>
<td>0.626</td>
</tr>
<tr>
<td>Positive Primes</td>
<td>-0.063</td>
<td>0.586</td>
</tr>
<tr>
<td>Average Appraisal</td>
<td>-0.250</td>
<td>0.532</td>
</tr>
</tbody>
</table>

Figure 3.1. Trustworthiness appraisals by prime condition and group.

To examine the dimensional association of trustworthiness appraisal and borderline features (i.e., number of criteria endorsed) across groups, I used a Spearman’s rho correlation
between borderline features and trustworthiness appraisals. This revealed an effect in the hypothesized direction for borderline features as measured by the SNAP-2 \((r_s = -.340, p = .002)\) and the SCID-II \((r_s = -.281, p = .013)\) such that more borderline features were associated with more negative trustworthiness appraisals. Table 3.5 includes correlations of borderline features and average trust appraisals. Additionally, Pearson correlations between normed measures of borderline features and average trustworthiness appraisals were significant. Contrary to expectation, BOR T scores were not correlated with average trustworthiness appraisals \((r = -.164, p = .155)\). However, the PAI affective instability (PAI-A) scale was correlated with average trust appraisals \((r = -.225, p = .050)\), as were SNAP-2 borderline personality T scores \((r = -.262, p = .021)\). Table 3.6 includes correlations of borderline scale T scores with average trust appraisals.

<table>
<thead>
<tr>
<th></th>
<th>SNAP-2 BPD Features</th>
<th>SCID-II BPD Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAP-2</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>SCID-II</td>
<td>.723**</td>
<td>--</td>
</tr>
<tr>
<td>Average Appraisals</td>
<td>-.340**</td>
<td>-.281*</td>
</tr>
</tbody>
</table>

*Note. Values in table are Spearman's rho correlations. ** \(p < .01\), * \(p < .05\) (two-tailed tests)
Hypothesis 2

Relative to the control group, the influence of negative primes on trustworthiness appraisals will be significantly greater for individuals in the borderline group.

I used two repeated-measures analysis of variance (ANOVA) models to test the influence of affective primes on trustworthiness ratings as a function of group. In the first model, I used a 2x3 ANOVA of appraisals with group (borderline vs. control) as the between-subjects factor and affective prime (positive, negative, neutral) as the within-subjects factor. In the second model, I used a 2x2 ANOVA of difference scores with group (borderline vs. control) as the between-subjects factor and difference scores (negative, positive), as the within-subjects factor.

For the first model, Mauchly’s test indicated violation of the sphericity assumption, $\chi^2(2) = .70.04$, $p < .01$. Results reported here correct conservatively for this violation using Greenhouse-Geisser adjustment to degrees of freedom. There were significant main effects of priming condition ($F(1.24, 93.06) = 24.74$, $p < .001$, partial $\eta^2 = .248$) and group ($F(1, 75) = 7.36$, $p = .008$, partial $\eta^2 = .089$) as well as a significant Prime x Group interaction ($F(1.24,$

\[ \text{Table 3.6. Correlations of borderline scale T scores with average trust appraisals.} \]

<table>
<thead>
<tr>
<th></th>
<th>PAI-BOR</th>
<th>PAI-A</th>
<th>PAI-I</th>
<th>PAI-N</th>
<th>PAI-S</th>
<th>SNAP-2 BOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAI-BOR</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAI-A</td>
<td>.900**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAI-I</td>
<td>.831**</td>
<td>.677**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAI-N</td>
<td>.828**</td>
<td>.633**</td>
<td>.640**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAI-S</td>
<td>.792**</td>
<td>.678**</td>
<td>.481**</td>
<td>.509**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNAP-2 BOR</td>
<td>.744**</td>
<td>.738**</td>
<td>.614**</td>
<td>.632**</td>
<td>.500**</td>
<td></td>
</tr>
<tr>
<td>Average Appraisals</td>
<td>-.164</td>
<td>-.225*</td>
<td>-.191</td>
<td>-.033</td>
<td>-.095</td>
<td>-.262*</td>
</tr>
</tbody>
</table>

Note. Values in table are Pearson correlations. PAI-BOR = Personality Assessment Inventory Borderline Scale. PAI-A, PAI-I, PAI-N, and PAI-S are subscales of the broader PAI-BOR scale (affective instability, identity problems, negative relationships, and self-harm, respectively).

** $p < .01$, * $p < .05$ (two-tailed tests)
This reveals a significant difference in the influence of affective priming on individuals in the borderline features group as compared to the control group. Table 3.4 shows differences between groups after each priming condition. Differences in appraisals between groups were significant in all priming conditions. Results are illustrated in Figures 3.1 and 3.2.

ANOVA results for the second model revealed a significant effect of difference score condition (negative vs. positive), $F(1, 75) = 29.35, p < .001, \text{partial } \eta^2 = .281$. In this model there was no main effect of group, $F(1, 75) = 2.46, p = .121, \text{partial } \eta^2 = .032$. The interaction of Prime x Group was significant, $F(1, 75) = 5.07, p = .027, \text{partial } \eta^2 = .063$. Independent samples t tests revealed that relative to the control group, appraisals made by individuals in the borderline features group were more influenced by negative affective primes relative to neutral affective primes ($p = .015$, one-tailed). Groups did not differ on positive difference scores. Table 3.7 shows the relevant means and results are further illustrated in Figures 3.3 and 3.4.

Table 3.7. Trust appraisal difference score averages by group.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Borderline</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Negative Difference</td>
<td>-0.2583</td>
<td>0.449</td>
</tr>
<tr>
<td>Positive Difference</td>
<td>0.1499</td>
<td>0.243</td>
</tr>
</tbody>
</table>

$t$  | $p$  | $R^2$  |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.22</td>
<td>$0.015$</td>
<td>0.062</td>
</tr>
<tr>
<td>1.20</td>
<td>0.12</td>
<td>0.019</td>
</tr>
</tbody>
</table>

*Note.* Significance values reflect one-tailed tests.
Figure 3.2. Trustworthiness appraisals by prime condition and group.

Figure 3.3. Negative and positive difference scores by group. These scores are an index of the influence of affective primes. Negative difference scores result from the subtraction of average ratings made after neutral primes from average ratings made after negative primes. The same procedure yields positive difference scores. Individuals in the borderline group showed a significantly greater priming effect: the influence of negative primes was significantly greater than for controls. Groups did not differ on positive difference scores.
I also expected that priming effect would be associated with number of borderline criteria endorsed, regardless of group. This was confirmed; borderline criteria as assessed by both the SNAP-2 and the SCID-II were correlated with negative difference scores. Endorsement of more criteria predicted lower difference scores (note that a difference score of 0 = no priming effect, and scores indicate greater effect of priming as they decrease from 0). Borderline criteria were also associated with raw trustworthiness ratings in all conditions: the more a person endorsed BPD symptoms, the lower their trustworthiness appraisals tended to be. Table 3.8 provides zero-order Spearman correlations among measures of borderline criteria, difference scores, and raw appraisals.
Hypothesis 3

Affective instability will predict priming effects.

Across groups I predicted that affective instability would be associated with trustworthiness ratings and priming effects. Zero-order correlations between the PAI borderline scale and subscales with difference scores and raw trustworthiness appraisals are shown in Table 3.9. Consistent with hypotheses, the affective instability subscale (PAI-A) was most stably related to trustworthiness ratings and priming effects. Scores on this scale were associated with greater negative priming effects as well as raw trustworthiness ratings following negative and neutral, but not positive, primes. Somewhat surprisingly, other subscales of the PAI were largely unrelated to any trustworthiness ratings or priming effects. One exception is that the identity problems subscale was associated with more negative appraisals made after negative primes. Although though these are correlational data and causal claims cannot be fully supported, the possibility that affective instability may be the specific dysfunction that drives biased trustworthiness appraisals warrants further consideration.
Hypothesis 4

Rejection sensitivity will mediate the relationship between borderline features and priming effects.

Rejection sensitivity was assessed using two measures: the Adult Rejection Sensitivity Questionnaire (ARSQ) and the McLean Assessment of Rejection Sensitivity (MARS). Both measures are described in greater detail in Chapter 2. Borderline features were related to both measures of rejection sensitivity. This is true for both SNAP-2 and SCID-II measurements of borderline features. Magnitudes of correlations, as included in Table 3.10, were largely similar across measures. Consistent with results from the correlational analysis, the borderline features group scored higher on both measures of rejection sensitivity than the control group. Group statistics are included in Table 3.11. Results were largely consistent across the two rejection sensitivity measures. For simplicity, I used only the ARSQ in further analyses because this is currently the most commonly used and best validated measure of rejection sensitivity in the BPD literature.

<table>
<thead>
<tr>
<th>Table 3.9. Pearson correlations of symptom groups with difference scores and raw trustworthiness ratings.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Negative Difference Score</td>
</tr>
<tr>
<td>Positive Difference Score</td>
</tr>
<tr>
<td>Raw Trustworthiness Ratings</td>
</tr>
<tr>
<td>Negative Prime</td>
</tr>
<tr>
<td>Neutral Prime</td>
</tr>
<tr>
<td>Positive Prime</td>
</tr>
</tbody>
</table>

*Note.* Correlations are with PAI T scores for each scale.

** **p < .01, * p < .05 (one-tailed tests)
Because groups did not differ on positive priming effects, mediation analyses were limited to negative priming (using negative difference scores). To test whether rejection sensitivity mediated the relationship between borderline features and negative priming effects, I regressed ARSQ scores on borderline features (as measured using the SNAP-2). This revealed a significant contribution of borderline features to rejection sensitivity scores ($\beta = .476$, $p < .001$).

Borderline features were then included as a predictor variable with negative difference scores as the outcome variable. This revealed a significant trend for borderline features to predict negative priming effect ($\beta = - .199$, $p = .083$). This effect is significant if evaluated using a one-tailed test, which is appropriate given the directional \textit{a priori} hypothesis. However, rejection sensitivity did not account for a significant portion of the variance in negative difference scores ($\beta = - .051$, $p = .

<table>
<thead>
<tr>
<th>Measure</th>
<th>SNAP-2 BPD Features</th>
<th>SCID-II BPD Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARSQ</td>
<td>0.51**</td>
<td>0.40**</td>
</tr>
<tr>
<td>Rejection Mechanism</td>
<td>0.51**</td>
<td>0.53**</td>
</tr>
</tbody>
</table>

\textit{Note.} Values in table are Spearman's rho correlations.

** $p < .01$, * $p < .05$ (two-tailed tests)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Borderline</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>ARSQ</td>
<td>11.29</td>
<td>4.90</td>
</tr>
<tr>
<td>Rejection Mechanism</td>
<td>7.70</td>
<td>5.45</td>
</tr>
</tbody>
</table>

\textit{Note.} Significance values reflect two-tailed tests.
.657) and was not a mediator of the trending association of borderline features and negative
priming effects.

I also tested whether rejection sensitivity mediated the relationship between borderline
features and raw average trustworthiness appraisals. Although this was not a primary hypothesis,
I did predict that these results would be consistent with past work that has shown full mediation
by rejection sensitivity in individuals with subclinical borderline features (Miano et al., 2013).
The first step in this analysis was the same as above: I confirmed that borderline features
accounted for a significant portion of the variance in rejection sensitivity scores ($\beta = .476, p < .001$). I then used borderline features as a predictive variable and average trustworthiness
ratings/appraisals as the outcome variable. This revealed that borderline features made a
significant contribution to average trustworthiness appraisals ($\beta = .263, p = .021$). However,
inconsistent with Miano et al.’s (2013) findings, rejection sensitivity did not account for a
significant portion of the variance in average trustworthiness appraisals ($\beta = -.056, p = .631$). The
same is true using either measure of rejection sensitivity ($\beta = -.144, p = .212$ for the MARS).
This indicates that although rejection sensitivity and borderline features are associated in this
sample, rejection sensitivity is not related to trustworthiness appraisals and thus did not mediate
the relationship between borderline features and these ratings.

Hypothesis 5

*The experience of childhood trauma will moderate the relationship between borderline
features and trustworthiness appraisals.*

Specifically, I expected that the emotional abuse and emotional neglect subscales of the
Childhood Trauma Questionnaire (CTQ; Berenstein & Fink, 1998) would moderate the
relationship of borderline group membership and trustworthiness appraisals. I expected this for two reasons: (1) emotional abuse and neglect are more “subtle” forms of early aversive childhood experiences that are consistently linked to borderline pathology (e.g., Linehan, 1993); (2) these were the only two subscales to differentiate the borderline and control groups (see Chapter 2 for measure description and group differences). Zero-order correlations between aspects of childhood trauma, average trustworthiness appraisals, and priming effects are summarized in Table 3.12.

<table>
<thead>
<tr>
<th></th>
<th>Emotional Abuse</th>
<th>Emotional Neglect</th>
<th>Physical Abuse</th>
<th>Physical Neglect</th>
<th>Sexual Abuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Average Trustworthiness Ratings</td>
<td>-0.13</td>
<td>-0.23*</td>
<td>-0.11</td>
<td>-0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>Negative Difference Score</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.02</td>
<td>-0.09</td>
<td>0.00</td>
</tr>
<tr>
<td>Positive Difference Score</td>
<td>-0.07</td>
<td>-0.03</td>
<td>-0.09</td>
<td>0.10</td>
<td>-0.08</td>
</tr>
</tbody>
</table>

*Note. Values in table are Pearson correlations.** p < .01, * p < .05 (two-tailed tests)

To examine whether childhood trauma moderated the relationship between borderline features/group membership and trustworthiness appraisals, I conducted a series of moderation analyses using linear regression. Dependent variables of interest were average trustworthiness appraisals and priming effects (negative and positive difference scores). Each dependent variable was examined in turn. At the first step of each moderation analysis, group (borderline vs. control) and childhood trauma (a dimensional variable representing one subscale of the CTQ at a
time) were entered. In the second step, the interaction term of group and trauma was entered. If this interaction term did not yield a significant F change, the analysis stopped.

**Emotional abuse.** Average trustworthiness ratings were included as the first dependent variable. Borderline group membership and CTQ emotional abuse were entered in step 1. This model was significant, $F(2, 74) = 3.63, p = .031$. A model with group membership, CTQ emotional abuse, and the interaction of group and emotional abuse was also significant, $F(2, 74) = 3.08, p = .033$. However, this second model did not yield a significant F change from model 1, indicating that the interaction term could not explain any additional variance above and beyond group membership and emotional abuse scores, $\Delta R^2 = .023, \Delta F(1, 73) = 1.89, p = .173$.

Additionally, the main effect of emotional abuse was not significant in model 1 ($\beta = -.001, p = .996$). There was a significant effect of borderline group in model 1 ($\beta = -.3, p = .017$). These results indicate that emotional abuse did not moderate the relation of borderline traits and average trustworthiness ratings.

I repeated this analysis for negative difference scores (an index of the impact of negative primes on trustworthiness appraisals). Step 1 of this model was not significant, $F(2, 74) = 2.48, p = .091$. There was again no main effect of emotional abuse ($\beta = -.04, p = .78$), and again a significant main effect of group membership emerged ($\beta = -.26, p = .037$). There was no significant change from model 1 to model 2 with the interaction term included, $\Delta R^2 = .003, \Delta F(1, 73) = .253, p = .616$. These results indicate that emotional abuse did not moderate the relation of borderline traits and negative priming effects.

I repeated the analysis once more for positive difference scores (an index of the impact of positive primes on trustworthiness appraisals). Step 1 of this model was not significant, $F(2, 74) = 1.47, p = .24$. There was again no main effect of emotional abuse ($\beta = -.15, p = .23$), and this
time there was not a significant main effect of group membership ($\beta = -.20, p = .11$). This is consistent with findings described above in which groups did not differ on the influence of positive primes. There was no significant change from model 1 to model 2 with the interaction term included, $\Delta R^2 = .032$, $\Delta F(1, 73) = 2.49, p = .119$. These results indicate that emotional abuse did not moderate the nonsignificant relation of borderline traits and positive priming effects.

**Emotional neglect.** Average trustworthiness ratings were included as the first dependent variable. Borderline group membership and CTQ emotional neglect were entered in step 1. This model was significant, $F(2, 74) = 4.1, p = .02$. A model with group membership, CTQ emotional neglect, and the interaction of group and emotional neglect was also significant, $F(3, 73) = 2.93, p = .039$. However, this second model did not yield a significant F change from model 1, indicating that the interaction term could not explain any additional variance above and beyond group membership and emotional neglect scores, $\Delta R^2 = .008$, $\Delta F(1, 73) = .619, p = .434$. Additionally, the main effect of emotional neglect was not significant in model 1 ($\beta = -.11, p = .36$). There was a significant effect of borderline group in model 1 ($\beta = -.25, p = .048$). These results indicate that emotional neglect did not moderate the relation of borderline traits and average trustworthiness ratings.

I repeated this analysis for negative difference scores (an index of the impact of negative primes on trustworthiness appraisals). Step 1 of this model was not significant, $F(2, 74) = 2.49, p = .09$. There was again no main effect of emotional neglect ($\beta = -.04, p = .76$), and again a significant main effect of group membership emerged ($\beta = -.26, p = .038$). There was no significant change from model 1 to model 2 with the interaction term included, $\Delta R^2 = .003$, 

73
$\Delta F(1, 73) = .226, p = .636$. These results indicate that emotional abuse did not moderate the relation of borderline traits and negative priming effects.

I repeated the analysis a final time for positive difference scores (an index of the impact of positive primes on trustworthiness appraisals). Step 1 of this model was not significant, $F(2, 74) = 1.16, p = .32$. There was again no main effect of emotional neglect ($\beta = -.12, p = .35$), and this time there was not a significant main effect of group membership ($\beta = .19, p = .14$). This is consistent with findings described above in which groups did not differ on the influence of positive primes. There was no significant change from model 1 to model 2 with the interaction term included, $\Delta R^2 = .009, \Delta F(1, 73) = .657, p = .42$. This indicates that emotional neglect did not moderate the nonsignificant relation between borderline traits and positive priming effect.

**Discussion**

I examined whether affective primes exert more influence over trustworthiness appraisals for individuals with borderline features relative to control participants, as well as how rejection sensitivity, childhood trauma, and state/trait mood variables interact with this response to affective primes. Several main findings resulted: (1) in agreement with a growing body of empirical research, individuals in the borderline group showed a general bias to rate faces as less trustworthy, regardless of affective primes; (2) in extension of previous work, results here indicate that negative affective primes exert a greater influence on trustworthiness ratings for individuals with borderline features relative to controls; (3) rejection sensitivity does not mediate the association of borderline features and trustworthiness appraisals in this sample; (4) the affective instability component of borderline pathology seems to relate uniquely to biased trust processing; and (5) the experience of childhood trauma does not moderate the association of borderline features and biased trustworthiness appraisals.
Individuals with borderline personality disorder or borderline features show a robust bias to rate others as less trustworthy than individuals without borderline features. This bias has now been demonstrated in multiple studies using a variety of methods and stimuli. The current study is the latest to confirm a trustworthiness bias. Average trustworthiness ratings, collapsed across affective priming condition, were significantly lower for individuals in the borderline group relative to the control group. Rapid impression formation for individuals with borderline personality disorder is biased such that others are viewed as less trustworthy. This may be construed as an interpretive bias—in the absence of clear information about the trustworthiness of others, individuals with borderline features have a bias to interpret ambiguity in favor of rating others as less trustworthy. This bias is clear, but its underlying mechanisms and etiology are not.

One potential contributor to a trustworthiness bias is affective information. This is the first study to show that individuals with borderline features are more greatly influenced by negative affective information when making trustworthiness appraisals than are individuals without borderline features. The influence of negative affective primes was not attributable to state changes in anxiety, positive mood states, or negative mood states. It was also not related to trait anxiety or social anxiety. The greater influence of negative affective priming on individuals in the borderline group cannot be attributed to a general priming effect. Groups did not differ in the degree to which they were influenced by positive affective primes. Thus it appears that individuals with borderline features are differentially and specifically influenced by negative affective information. This is not surprising in light of a great wealth of literature indicating that borderline personality is associated with higher levels of negative affect. These findings suggest that while a negative bias in trustworthiness ratings is relatively stable for individuals with borderline features, this bias may be enhanced in the context of negative affective information.
One potential mechanism underlying the association of borderline features and biased trustworthiness appraisal is rejection sensitivity. Past work has shown that rejection sensitivity mediates the association of borderline features and trust appraisals (Miano et al., 2013). However, the results of the current study do not support this link: rejection sensitivity (assessed with the same measure as Miano et al., 2013) was not related to trustworthiness appraisals or to priming effects. Future research should seek to reconcile this somewhat puzzling result. It is possible that the sample used by Miano et al. (2013) was more severe. However, this is unlikely given that Miano et al. also examined individuals with borderline features who did not necessarily meet full diagnostic criteria for BPD. Indeed, the reverse is more likely true, as Miano et al. used an undergraduate sample (mean age = 19.8). However, rejection sensitivity was not related to trust appraisals even in the current control group. Overall, the current study raises questions about the plausibility of rejection sensitivity as a mechanism for biased trust appraisal.

As predicted, the impact of negative affective primes was uniquely predicted by the affective instability subscale of the broader PAI Borderline scale. This subscale was associated with negative priming effects (negative difference scores), as well as appraisals made after negative primes and neutral primes. It was not associated with ratings made after positive primes, providing further evidence that positive and negative emotional processing in BPD may rely on separable mechanisms that influence and are influenced by differentiated systems. The PAI identity disturbance subscale was also related to appraisals made after negative primes, but not to negative difference scores/negative priming effects. Overall, it seems that the exaggerated influence of negative affective information that was indexed on the behavioral task used here validly captured affective instability as reported on the PAI. Individuals who themselves reported greater difficulty regulating affect showed behavioral evidence indicative of the same.
Interestingly, no trustworthiness ratings (regardless of priming condition) were related to the negative relationships subscale of the PAI. One overarching goal of the current study is to better characterize the impact of biased social cognition on measures of social functioning (see Chapters 5 and 6). It appears that the two may be dissociable aspects of borderline personality.

Like rejection sensitivity, childhood trauma also did not emerge as a significant predictor of biased trustworthiness appraisals. Childhood trauma also did not moderate the relation of borderline features and biased trustworthiness ratings. This is somewhat surprising given past work that suggests a direct link between early aversive experiences and trust appraisal. One possible explanation for the findings is the seemingly high rate of minimization/denial evidenced in the control group (see Chapter 2). If childhood trauma was not accurately or consistently reported for the entirety of the sample, the results here may not be fully interpretable.

Alternatively, it may be the case that early childhood trauma may influence trust processing in differential ways—other work has suggested that betrayal traumas may produce a positive trustworthiness bias in some individuals such that they are too trusting of others (Gobin & Freyd, 2009; Zurbriggen & Freyd, 2004). Further work is needed to better understand what causes these two divergent consequences following childhood trauma. Additionally, the current sample may have also reflected a lower base rate of childhood abuse than other samples, which may truncate variance and make the results difficult to interpret. The borderline group did not differ from the control group on several subscales of the Childhood Trauma Questionnaire, including physical abuse, sexual abuse, and physical neglect. Finally, it may be the case that childhood trauma influences other aspects of social cognition and trustworthiness appraisals.

Future work is needed to understand why negative affective information seems to exert undue influence on trustworthiness appraisal in borderline pathology, as well as whether this
influence extends to other aspects of social cognition. One possible mechanism for undue affective influence is problems with cognitive control of this information. Dysfunctional regulatory control over an over-reactive amygdala may explain the affective priming effect seen here. Work that integrates neuroimaging with this type of behavioral task may prove fruitful for better understanding social cognitive biases in BPD. Given that amygdala habituation may improve emotion regulation over the course of dialectical behavior therapy (Goodman et al., 2014), it may also be fruitful to examine how social cognitive biases and the influence of affective information on these biases changes over the course of treatment. BPD has also been associated with deficits in executive functions, of which cognitive control over emotional information is one sub-type (Gvirts et al., 2012; Ruocco et al., 2012). The association of these sometimes subtle cognitive differences and social cognition is a suggested target for future work.

When considered in summary, the results of the current study suggest that trustworthiness appraisal is one aspect of social cognition that is reliably biased in borderline pathology. This bias can be augmented with negative affective information. This in turn suggests a connection between borderline pathology and difficulty exerting cognitive control over negative affective information. Trustworthiness appraisal bias is not explained by rejection sensitivity, trait anxiety, social anxiety, mood changes, or childhood trauma, and seems uniquely related to affective instability. Although this type of appraisal falls squarely within the realm of social cognition, the correlational data from this study suggest it is not related to self-reported difficulty in relationships. This topic is explored in greater detail in Chapter 6.
Chapter 4: Face Recognition and Trustworthiness Appraisal
Introduction

In the preceding chapter the borderline group demonstrated a bias to rate faces as untrustworthy, relative to the control group. Participants in the borderline group also showed greater susceptibility to the influence of negative affective primes when making trustworthiness judgments: they made significantly lower trustworthiness appraisals following negative emotional images. In other words, the difference in judgments made following neutral versus negative emotional information was significantly greater for the borderline group compared to the control group. While these findings are important in their own right, it is important to consider whether trustworthiness biases show stability over time. It is also important to consider how the perception of others as less trustworthy might be associated with one’s ability to recognize whether they have seen that person before as well as their certainty in making a claim that a person is either someone they recognize or someone new. These issues are examined in the current chapter.

Appraisal Stability

Past work in non-clinical populations has shown that trait judgments are relatively stable over short durations (Todorov et al., 2005). Judgments made very quickly based on facial information also endure over longer exposure to faces (Willis & Todorov, 2006). To date, however, no work has examined whether trustworthiness bias shows stability across a delay for individuals with borderline personality features or disorder. Here it was predicted that the borderline group would continue to show a trustworthiness appraisal bias even after a short delay. Additionally, for the entire sample it was expected that average trustworthiness appraisal made during the affective priming task would significantly predict average trustworthiness appraisal made following a delay, during a face recognition task.
Recognition and Trustworthiness

Memory biases and difficulties have been identified in BPD. Meta-analysis shows that individuals with BPD have diminished verbal and nonverbal memory relative to healthy controls (Ruocco, 2005). Research on autobiographical memory biases in BPD has been less clear: some evidence indicates an overgeneral memory bias in BPD (e.g., Jones et al., 1999) and some indicates no difference in autobiographical memory between controls and individuals with BPD (Arntz, Meeren, & Wessell, 2002; Renneberg, Theobald, Nobs, & Weisbrod, 2005). Tests of autobiographical memory examine the extent to which individuals are able to recall specific versus overgeneralized memories. A tendency to recall overgeneralized memories has a potentially protective function; these memories may exclude threatening or emotional details that characterize specific memories (see Renneberg et al., 2005). Although it remains unclear whether individuals with BPD have biases in autobiographical memory, there is evidence that these individuals have a bias to recall a greater number of negative memories on autobiographical memory tests, without the protracted reaction time that is characteristic of depression (Renneberg et al., 2005). In other words, individuals with BPD quickly access negative memories.

Memory biases in BPD, including the tendency to access greater numbers of negative memories relative to positive or neutral memories, are often context-dependent. For example, in some studies memory is only impaired in BPD in the context of negative emotional information (Menesbach et al., 2009). In other words, compared to neutral or positive information, negative information differentially interferes with encoding and retrieval processes. This is often referred to as negative affective interference. A recent fMRI investigation by Soloff and colleagues (2015) of neural processing during cognitive tasks showed that negative affect interfered with a
range of cognitive control processes upon which memory relies. During an episodic memory task, individuals with BPD showed decreased activation in the hippocampus, anterior cingulate cortex (ACC), precuneus, and dorsal lateral prefrontal cortex (DLPFC) relative to control participants. When asked to recognize previously viewed affective images, individuals with BPD showed decreased activation in the ACC, DLPFC, and hippocampus for negative versus positive images. Overall, the effect of negative affective information on cognitive processes was greater for individuals with BPD relative to controls. Growing evidence now suggests that negative information interferes with cognitive processes to a greater extent in individuals with BPD than controls—this may be the essence of emotion dysregulation in BPD (Soloff et al., 2015).

Consistent with Soloff et al.’s (2015) finding that individuals with BPD recruited less aid from the DLPFC in the context of negative affective information, BPD is associated with more difficulty inhibiting negative emotional information in the service of meeting task demands (e.g., Arntz Appels, & Sieswerda, 2000; Domes et al., 2006). When actively attempting not to give attention to negative information, individuals with BPD seem to be more influenced or pulled by this information in such a way that results in performance impairment.

As described below, participants were unexpectedly asked after a delay to remember which faces they had seen during the affective priming task. Some faces were faces they had seen before, and some faces were new. They were asked to identify whether they had seen the face before (yes/no). Responses were categorized as outlined in Table 4.1. Correct responses included hits (i.e., a participant responded that he had seen the face before and this was correct) or true negatives (i.e., the participant responded that she had not seen the face before and this was correct). Incorrect responses included false positives (i.e., a participant responded that he had seen the face before and this was incorrect—he falsely reported having seen it) and misses.
(i.e., the participant responded that she had not seen the face before and this was incorrect—she missed a face that was present before).

<table>
<thead>
<tr>
<th>New (Never Seen)</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect: False Positive</td>
<td>Correct: True Negative</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Old (Seen Before)</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct: Hit</td>
<td>Incorrect: Miss</td>
<td></td>
</tr>
</tbody>
</table>

The affective priming task described in Chapter 3 and the recognition task used here included facial stimuli. These stimuli themselves convey affective information. Facial cues of untrustworthiness, for example, are inherently signals of social danger or threat. Untrustworthy individuals are more likely to reject, betray, or harm others. Judgments about traits like trustworthiness are made rapidly (i.e., in as little as 100ms) and subconsciously (Willis & Todorov, 2006). This makes faces regarded as untrustworthy themselves a type of negative affective information that might interfere with accurate recognition of the faces. Given that individuals with borderline features consistently rate faces as less trustworthy compared to controls, faces should be more difficult to recognize accurately. Thus it was expected that compared to control participants, individuals with borderline features would make more errors overall (misses and false positives) on a recognition task. Additionally, it was expected that overall trustworthiness ratings would be lower for faces the borderline group misremembered. In other words, it was expected that recognition accuracy in the borderline group would be
differentially impacted as function of the degree to which participants in this group perceived faces as untrustworthy in-the-moment, during the recognition task. Faces misremembered would be those that the borderline group perceived as less trustworthy in-the-moment. This effect was not expected in the control group: in-the-moment perceptions of trustworthiness were not expected to differ for accurately identified versus misremembered faces in this group.

**Certainty in Recognition Judgments**

Individuals with BPD have been shown to differ from individuals with schizophrenia in that they *over*-mentalize social judgments rather than *under*-mentalize (Andreou et al., 2015). In other words, they think more actively about and make more complicated attributions about the mental states of others. On the trustworthiness task, this might manifest as more extreme ratings of trustworthiness (either high or low), as extreme ratings can be considered a marker of over-active or complex mental state attributions. Over-mentalization of social judgments is associated with overconfidence in errors, such that even when making an error individuals who over-mentalize are more confident that they are correct (Andreou et al., 2015). Individuals in the borderline group showed a bias to rate faces as more extremely untrustworthy compared to the control group. In terms of mentalizing, this signals that they may be making overactive attributions about the mental states of the individuals depicted. Accordingly, I expected that individuals with borderline features would show greater confidence in their recognition judgments overall relative to controls. Because the borderline group consistently rated faces as less trustworthy in Chapter 3, and this biased attribution could contribute to over-mentalization and difficulty monitoring the source of information (i.e., more errors), it was also expected that relative to the control group, the borderline group would be more confident or certain about their incorrect responses.
In summary, the following specific hypotheses were made: (1) average trustworthiness appraisals would again be lower in the borderline group compared to the control group; (2) average trustworthiness appraisals made during the earlier affective priming task would predict average trustworthiness appraisals made during the recognition task; (3) the borderline group would show lower overall recognition accuracy than the control group; (4) trustworthiness ratings for incorrect responses versus correct responses would be significantly lower in the borderline group, but the control group would not demonstrate this effect; and (5) the borderline group would express greater certainty about incorrect responses compared to the control group.

**Method**

Participants first completed the affective priming task described in Chapter 3. During this task they rated the trustworthiness of 50 faces. Participants made trustworthiness ratings three times for each face, with one rating made following each type of prime (negative, neutral, or positive emotional images). Each face was presented for 7 seconds on 3 occasions for a total stimulus exposure of 21 seconds for each face. Participants were asked to study each face when making their trustworthiness judgments, and were told that the speed with which they made ratings would not dictate how quickly the task progressed.

Following this rating task, participants completed the PANAS and the STAI (see Chapters 2 and 3 for descriptions of these measures). They then completed items on a basic social functioning measure for a short delay (11-21 minutes, $M = 13.30$, $SD = 1.95$), before completing the memory task. The questions they completed during this window of time were simple and straightforward (e.g., are you a student?), with low likelihood of emotional impact. They consisted of questions about the person’s typical work, educational pursuits (as applicable), and housework. After the delay, participants were instructed that they would view a series of
faces, this time with no scenes before the faces. Participants were told that some of the faces were faces they had seen during the earlier task, but that some of the faces were new. They were told that they would be asked whether they had seen the face before, and that sometimes the correct answer would be “yes,” and sometimes the correct answer would be “no.” They would then be asked to rate how certain they were about this judgment, as well as to rate the trustworthiness of each face.

The task was computerized using the same program (MatLab) and had the same overall appearance as the affective priming task that participants had completed earlier. Stimuli included the 50 faces used during the affective priming task (25 female, 25 male), as well as 26 new faces (13 female, 13 male) from the same face set (which was originally used by Adolphs et al., 1998). Faces were presented in random order and duration of presentation was determined by participant response. Once the participant had answered all three questions, the task proceeded to the next face. The face appeared in the center of the screen with each question printed below. Exact question wordings were “Have you seen this face before?”, “How certain are you?” and “How trustworthy is this person?” Certainty and trustworthiness were assessed on a -3 to 3 scale, where -3 corresponded to “not certain at all” or “not trustworthy at all” and 3 corresponded to “completely certain” or “extremely trustworthy.” Yes/no recognition responses were classified based on accuracy as described earlier in Table 4.1.

Results

Memory Task Variables

For the overall sample, trustworthiness ratings made during the memory task were not associated with certainty ratings, $r(77) = \cdot003, p = .976$. In other words, participants were not more or less certain about the accuracy of their yes/no responses when they appraised faces as
trustworthy. There was a significant trend in the association of trustworthiness ratings and accuracy, \( r(77) = .213, p = .063 \). The new trustworthiness ratings made during the recognition task trended toward predicting accuracy—participants were more likely to respond correctly if they appraised a face as trustworthy in-the-moment. Similarly, a trend emerged in the association of certainty ratings with accuracy, \( r(77) = .199, p = .083 \). As might be expected, participants were more certain of their own accuracy when they made correct responses. Associations between variables in the memory task are presented in Table 4.2. “Time 2” in this table refers to ratings made during the recognition task as opposed to the trustworthiness ratings made during the earlier affective priming task. As listed in Table 4.3., the control group expressed greater certainty in their recognition judgments than the borderline group \( t(75) = 3.5, p = .001 \).

<table>
<thead>
<tr>
<th>Table 4.2. Pearson correlations of recognition task variables.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time 2 Recognition Task</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>Trustworthiness</td>
</tr>
<tr>
<td>Certainty</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
</tbody>
</table>

** \( p < .01 \), * \( p < .05 \), ^p < .1 \) (one-tailed tests)

**Hypothesis 1**

_Average trustworthiness appraisals made during the recognition task will be lower in the borderline group compared to the control group._

Consistent with findings from Chapter 3, when asked to assess the trustworthiness of faces on the second occasion during the recognition task, the borderline group again rated faces as significantly less trustworthy than the control group, \( t(75) = 2.47, p = .016 \). Table 4.3 includes
additional information, including group means. The general trustworthiness bias demonstrated in Chapter 3 remained present after a short delay.

### Hypothesis 2

*Average trustworthiness appraisals made during the earlier affective priming task will predict average trustworthiness appraisals made during the recognition task.*

For the entire sample, average trustworthiness appraisals made during the affective priming task significantly and strongly predicted average trustworthiness appraisals made during the recognition task after the delay, $r(75) = .930, p < .001$. This was also true when considering the control ($r(45) = .943, p < .001$) and borderline groups ($r(28) = .877, p < .001$) in isolation. This can be considered a rough measure of the test-retest reliability of trustworthiness appraisals across a short delay.

### Hypothesis 3

*The borderline group will show lower overall recognition accuracy than the control group*

Inconsistent with this prediction, overall accuracy rates did not differ by group, $t(75) = .681, p = .498$. Borderline and control groups correctly identified whether or not they had seen faces before at comparable rates. To determine whether groups differed across different types of
errors (false positives or misses), I used a repeated-measures analysis of variance (ANOVA). Specifically, I used a 2x2 ANOVA of response rate (i.e., number of responses that fit within that category) with error type (false positive vs. miss) as the within-subjects factor and group (borderline vs. control) as the between-subjects factor. False positive and miss rates were corrected as a function of overall opportunities to make each type of error (there were 50 opportunities to make a “miss” error and 26 opportunities to make a “false positive” error). There was a trending toward a main effect of error type \( (F(1, 75) = 3.37, p = 0.07, \text{partial } \eta^2 = 0.043) \)). Participants nearly made significantly more “miss” errors than “false positive” errors, \( t(76) = 1.90, p = 0.06 \). There was no main effect of group \( (F(1, 75) = 0.62, p = 0.43, \text{partial } \eta^2 = 0.01) \), consistent with the previous finding that groups did not differ on overall accuracy. The interaction of Error Type x Group was not significant \( (F(1, 75) = 0.001, p = 0.98, \text{partial } \eta^2 = 0.00) \)).

With regard to correct responses, groups did not differ significantly on rates of true negatives \( (t(75) = .735, p = .465) \) or hits \( (t(75) = .486, p = .628) \). Table 4.4 provides additional information, including group means. Results are illustrated in Figure 4.1. Attention to high accuracy rates and low base rates of false positive and misses is warranted.
Table 4.4. Group means and differences by on recognition categorization.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Borderline M</th>
<th>SD</th>
<th>Control M</th>
<th>SD</th>
<th>t</th>
<th>p</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy %</td>
<td>92.37</td>
<td>6.49</td>
<td>90.82</td>
<td>11.33</td>
<td>0.68</td>
<td>0.50</td>
<td>0.006</td>
</tr>
<tr>
<td>False Positives</td>
<td>1.40</td>
<td>2.43</td>
<td>1.79</td>
<td>2.14</td>
<td>0.74</td>
<td>0.47</td>
<td>0.007</td>
</tr>
<tr>
<td>True Negatives</td>
<td>24.60</td>
<td>2.43</td>
<td>24.21</td>
<td>2.13</td>
<td>0.74</td>
<td>0.47</td>
<td>0.007</td>
</tr>
<tr>
<td>Hits</td>
<td>45.60</td>
<td>3.79</td>
<td>44.81</td>
<td>8.38</td>
<td>0.49</td>
<td>0.63</td>
<td>0.003</td>
</tr>
<tr>
<td>Misses</td>
<td>4.40</td>
<td>3.79</td>
<td>5.19</td>
<td>8.38</td>
<td>0.49</td>
<td>0.63</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Note. Significance values reflect two-tailed tests.

Figure 4.1. Response rate by category and group
Hypothesis 4

*Trustworthiness ratings for incorrect responses versus correct responses will be significantly lower in the borderline group, but the control group will not demonstrate this effect.*

A repeated-measures ANOVA was used in which the within-subjects factor was defined as incorrect (false positives + misses) versus correct (hits + true negatives) responses and group (borderline vs. control) was included as the between-subjects factor. The dependent variable was trustworthiness appraisal. Average trustworthiness appraisals were calculated for each of the response types (correct vs. incorrect responses). Degrees of freedom in these analyses reflect the fact that not all participants made incorrect responses. Included in these analyses are 41/47 individuals in the control group and 29/30 individuals in the borderline group. There was no significant main effect of response type ($F(1, 68) = 1.28, p = 0.26, \text{partial } \eta^2 = 0.018$)). Overall, regardless of group, trustworthiness appraisals did not differ for incorrect versus correct responses. There was a significant main effect of group ($F(1, 68) = 11.03, p = 0.001, \text{partial } \eta^2 = 0.14$)), consistent with the tendency for individuals in the borderline group to rate faces as less trustworthy overall. The interaction of Response Type x Group was significant ($F(1, 68) = 4.68, p = 0.34, \text{partial } \eta^2 = 0.064$)). For correct responses, the borderline group rated faces as significantly less trustworthy than the control group, $t(68) = 2.42, p = .018$. The same was true of incorrect responses, $t(68) = 3.34, p = .001$. Trustworthiness ratings made by the control group did not vary across correct vs. incorrect responses, $t(40) = .937, p = .354$. However, the difference in trustworthiness ratings made by the borderline group across incorrect and correct responses is trending toward significant, $t(28) = 1.83, p = .078$. Faces to which this group responded incorrectly were concurrently rated as less trustworthy. Thus it appears that judgment accuracy was not related to trustworthiness in the control group, but that incorrect judgments for
the borderline group were typically made for faces rated as less trustworthy in-the-moment. Results are illustrated in Figure 4.2 and are largely consistent with hypothesis 4.

These results taken together suggest a general negativity bias in trustworthiness ratings for the borderline group. They also suggest that the momentary perception of faces as less trustworthy may contribute to higher rates of incorrect responses in general for individuals with borderline features. This is consistent with the notion that low trustworthiness conveys negative emotional information that may differentially interfere with cognitive processes, including memory retrieval, for individuals with borderline features. Group means for trustworthiness appraisals made during the recognition task are presented in Table 4.5., separated by response type.
Hypothesis 5

The borderline group will express greater certainty about incorrect responses than the control group.

As reported earlier, the borderline group reported less confidence than the control group about their ability to correctly recognize faces. In other words, after indicating if they had seen faces before (yes or no), they expressed less certainty that they were correct than the control group. However, it was expected that individuals with borderline features would show greater confidence in their incorrect recognition judgments specifically, relative to controls. This is because past work has shown that individuals with BPD over-mentalize on social judgment tasks, a process that supports increased confidence in their errors.

A repeated-measures ANOVA was used for which the within-subjects factor was defined as incorrect (false positives + misses) versus correct (hits + true negatives) responses. Average certainty ratings were calculated for each of these variables (correct vs. incorrect responses) and used as the dependent variable. Group (borderline vs. control) was entered as the between-groups factor. Degrees of freedom in these analyses reflect the fact that not all participants made incorrect responses. Included in these analyses are 41/47 individuals in the control group and 29/30 individuals in the borderline group. There was a significant main effect of response type

<table>
<thead>
<tr>
<th>Category</th>
<th>Borderline</th>
<th>Control</th>
<th>t</th>
<th>p</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>False positive</td>
<td>18 -0.90</td>
<td>31 0.43</td>
<td>2.85</td>
<td>0.01</td>
<td>0.147</td>
</tr>
<tr>
<td>True negative</td>
<td>30 -0.24</td>
<td>47 0.24</td>
<td>2.91</td>
<td>0.005</td>
<td>0.102</td>
</tr>
<tr>
<td>Hit</td>
<td>30 -0.15</td>
<td>46 0.22</td>
<td>2.21</td>
<td>0.030</td>
<td>0.062</td>
</tr>
<tr>
<td>Miss</td>
<td>25 -0.28</td>
<td>46 0.22</td>
<td>1.53</td>
<td>0.132</td>
<td>0.040</td>
</tr>
<tr>
<td>Correct</td>
<td>29 -0.21</td>
<td>41 0.20</td>
<td>2.42</td>
<td>0.018</td>
<td>0.079</td>
</tr>
<tr>
<td>Incorrect</td>
<td>29 -0.56</td>
<td>41 0.31</td>
<td>3.34</td>
<td>0.001</td>
<td>0.141</td>
</tr>
</tbody>
</table>

Table 4.5. Group means and differences on trustworthiness appraisal by response categorization.

Note: Significance values reflect two-tailed tests.
Overall, participants expressed greater certainty about correct responses relative to incorrect responses, $t(69) = 10.54, p < .001$. There was a significant main effect of group ($F(1, 68) = 8.38, p = 0.005$, partial $\eta^2 = 0.110$). The borderline group expressed less certainty than the control group overall, $t(68) = 2.89, p = .005$. Inconsistent with hypothesis 5, the interaction of Response Type x Group was not significant ($F(1, 68) = 2.38, p = 0.13$, partial $\eta^2 = .034$). For correct responses, the borderline group expressed significantly less certainty in their judgments than the control group, $t(68) = 3.11, p = .003$. The same was true of incorrect responses, $t(68) = 2.42, p = .018$. Certainty ratings made by the control group were significantly lower for incorrect responses than for correct responses, $t(40) = 6.92, p < .001$. This was also true of the borderline group, $t(28) = 8.47, p < .001$. Group certainty means across response types are included in Table 4.6 and depicted in Figure 4.3.

| Category     | Borderline | | | Control | | |
|--------------|------------|-----------|--------------|-----------|-----------|---------|---------|---------|---------|
|              | $n$ | $M$ | $SD$ | $n$ | $M$ | $SD$ | $t$ | $p$ | $R^2$ |
| False positive | 18 | 0.55 | 1.57 | 31 | 1.18 | 1.53 | 1.38 | 0.17 | 0.039 |
| True negative | 30 | 1.73 | 0.78 | 47 | 2.25 | 0.56 | 3.40 | 0.001 | 0.134 |
| Hit          | 30 | 2.42 | 0.42 | 46 | 2.68 | 0.29 | 3.08 | 0.003 | 0.114 |
| Miss         | 25 | 0.09 | 1.33 | 33 | 0.89 | 1.37 | 2.24 | 0.029 | 0.082 |
| Correct      | 29 | 2.10 | 0.52 | 41 | 2.45 | 0.40 | 2.42 | 0.018 | 1.240 |
| Incorrect    | 29 | 0.31 | 1.30 | 41 | 1.10 | 1.39 | 3.11 | 0.003 | 0.079 |

*Note.* Significance values reflect two-tailed tests.
Discussion

The current chapter considered whether trustworthiness appraisals show approximate test-retest reliability across a short delay, the influence of facial trustworthiness appraisals on memory for faces, accuracy of recognition judgments, and expression of certainty regarding memory judgments. A few main findings emerged: (1) negative trustworthiness appraisal bias in the borderline group relative to the control group was consistent even across a short delay; (2) trustworthiness ratings were largely consistent after a short delay for the overall sample—individuals who rated faces as less trustworthy overall during the affective priming task described in Chapter 3 also tended to rate faces as less trustworthy during the recognition task; (3) individuals in the borderline and control groups did not differ in their overall accuracy—they correctly identified new versus old faces at statistically equivalent rates; (4) the borderline group expressed significantly less certainty in their responses, regardless of whether those responses
were correct or incorrect; and (5) preliminary evidence suggests that the borderline group was differentially influenced by faces they rated as less trustworthy in-the-moment. In other words, they were more likely to make incorrect responses to faces they concurrently appraised as less trustworthy. Trustworthiness appraisals were not related to accuracy for the control group.

Individuals with borderline personality features consistently make lower trustworthiness appraisals. This has now been replicated across studies and over a short delay in the current study. Average trustworthiness ratings made after the affective priming task correlated highly for all participants with average trustworthiness appraisals made during the recognition task after a short delay. At both time points, the borderline group made significantly lower appraisals of trustworthiness. This adds to evidence of an association between trustworthiness bias and borderline personality. Notably, however, the delay period in this study was relatively short ($M = 13.30$ minutes). This is longer than typical short recall delays on neuropsychological tests (often 10 minutes), but shorter than typical long delay periods (typically 15-20 minutes). Further work can extend this initial evidence for test-retest reliability by implementing longer delay times.

Individuals with borderline features did not differ from the control group on overall accuracy. This is interesting in light of previous work that has indicated differences in memory function for individuals with BPD compared to controls. The lack of differences between groups observed here may be attributable to the relatively short delay time. It may also be attributable to the manner in which groups were defined. As described in Chapter 2, the borderline group was defined by endorsement of 3 or more BPD symptoms on the SNAP-2. Thus this group is best regarded as representing individuals with borderline features who did not necessarily meet full diagnostic criteria for BPD. Memory differences between this group of individuals with borderline features rather than full diagnosis and individuals in the control group may have been
attenuated relative to past studies that have compared individuals on the basis of diagnostic status. Another potential explanation for the lack of difference is the disparity in estimated IQ between the borderline and control groups. The borderline group in this study had a higher estimated IQ than the control group, and might reasonably be expected to perform better on tests of neurocognition.

Although the borderline group did not show a difference in accuracy of recognition relative to the control group, an important finding emerged with regard to trustworthiness appraisal and accuracy. The borderline group was more likely to make an incorrect response (false positive or miss) for faces they concurrently rated as less trustworthy in the memory task. In other words, incorrect responses were associated with lower in-the-moment trustworthiness ratings than correct responses for the borderline group. The control group did not show any similar association between trustworthiness and response accuracy. This finding provides further evidence that individuals with borderline features are more susceptible to negative emotional information. Cues of low trustworthiness (whether they are objectively present or perceived owing to response bias) are fundamentally a type of negative social information. Untrustworthy individuals are more likely to initiate betrayal, rejection, or other aversive social experiences. Evidence converges to show that negative emotional or social information differentially influences cognitive processes for individuals with BPD. The findings of the current study suggest that negative information also has a differential impact on cognitive processes in the context of a socially relevant task.

The finding that individuals in the borderline group made lower trustworthiness appraisals for faces to which they responded incorrectly may also be considered an example of a source monitoring error. Although source memory seems to be intact overall in BPD, lower
source memory is associated with greater hostility and suspiciousness for individuals with the disorder, even controlling for negative affect (Minzenberg et al., 2006). This means that individuals with BPD who are high on hostility and suspiciousness have greater difficulty monitoring the initial source of their information—whether that information was actually perceived or whether it was imagined. Further work may do well to examine how personal hostility and suspiciousness interact with social judgments to predict source monitoring errors. It may be the case that the perception of others as less trustworthy confuses some individuals with borderline features to believe that they have seen the face before.

Finally, individuals with borderline features expressed less certainty about their recognition judgments overall than individuals in the control group. In contrast to our expectations, these individuals did not express greater certainty than the control group about incorrect responses. This is seemingly inconsistent with ideas about hyper-mentalizing in BPD (e.g., Andreou et al., 2015). Perhaps promisingly for individuals with borderline pathology, this finding suggests some awareness of the difficulty of remembering this social information. Although individuals with borderline features may be more influenced by negative social information as they attempt to recall faces, they are also seemingly aware of their own fallibility.
Chapter 5: Measuring Social Functioning
Introduction

Social functioning is inherently difficult to assess. The nature of adaptive social functioning may vary across participants, and measures that rely on self-report are subject to standard biases associated with appraising and reporting upon one’s own behavior. Social functioning is typically measured either via self-report with cross-sectional, reflective measures, or with more resource-intensive methods, including experience-sampling and daily diaries.

Self-report measures of social functioning are subject to biased reporting and memory effects. Measuring social function via self-report in a disorder marked by general negativity presents an additional challenge. Borderline features are associated with negative appraisals of relational experiences (Barnow et al., 2009) and negative judgments even of positive social interactions (e.g., Bhatia et al., 2013). Thus individuals with borderline traits may report their social experiences as more negative overall than individuals low on borderline traits regardless of the actual content of those experiences. This does not negate the fact that an individual’s own perception of his/her social functioning is relevant and worthy of examination. In fact, despite the limitations of self-report, this type of strategy has been shown to have high clinical utility in populations of individuals with borderline traits, including sensitivity to change in distress and social functioning over a period of 6 months (Stepp et al., 2011). However, the potential remains for perceptual biases to distort our understanding of social functioning in BPD.

There are a wide range of self-report measures that can be used to assess social functioning. Problematically, many of the scales in general use concern only very narrow and specific domains of functioning or when examined on an item-level, appear to lack face validity. A number of interview measures also exist, and these may help eliminate some self-report bias. However, they are often resource intensive, as are newer methods that use experience sampling.
and/or daily diary recordings. In the sections that follow, some of the most important measures that have been developed to date are described and discussed.

**Social Function Questionnaire**

The Social Function Questionnaire (SFQ; Tyrer et al., 2005) is a brief, 8-item self-report scale developed from a semi-structured interview. Despite its brevity and the fact that it does not capture social functioning across a broad range of domains, the scale is psychometrically sound. However, it was not developed specifically in the context of research on BPD and item-level inspection suggests that some of the specific questions are closely related to symptoms of BPD (e.g., “I feel lonely and isolated from other people”). Thus the scale may be more of a measure of borderline pathology than of general social functioning.

**Inventory of Interpersonal Problems**

The Inventory of Interpersonal Problems Personality Disorders 25 (IIP-PD-25; Kim & Pilkonis, 1999) is a commonly used measure of social function in BPD. While this measure has strong psychometric properties, an examination of the items (e.g., “I am too envious of other people”) reveals that they may in some cases be a better measure of negative self-concept or general negativity than actual functioning. Some of these items are particularly problematic in light of the negative self-concept that is often found in BPD, as they include self-judgments (e.g., “I am too envious…”) that are likely to reflect bias in addition to reality. Other version of the IIP are available (e.g., the Inventory of Interpersonal Problems Circumplex; IIP-C; Horowitz et al., 2000) and are influenced by similar problems.

**Longitudinal Interval Follow-Up Evaluation**

The Longitudinal Interval Follow-Up Evaluation (LIFE; Keller et al., 1987) is the measure of choice for one of the largest ongoing longitudinal studies of BPD: the Collaborative
Longitudinal Personality Disorder Study (CLPS; see Gunderson et al., 2011). This evaluation includes a clinician-administered interview and assesses role functioning in several domains: employment, spousal relationships, parenting, friendships, and recreation. It also assesses satisfaction with role functioning. The evaluation is particularly suited to longitudinal evaluation, as it provides a simple system for determining how role functioning changes over time. However, the initial assessment is resource intensive and thus unlikely to gain traction outside of large, well-funded studies. The face-to-face interview requires 45-60 minutes. With coding included, the measure requires as long as 2-4 hours of researcher time for each participant. Moreover, although this measure is helpful in avoiding some self-report biases, it still relies upon participant report.

**Social Adjustment Scale**

Like the LIFE, the Social Adjustment Scale (SAS; Weissman et al., 1976) assesses functioning across several domains. However, unlike the LIFE, the SAS is a self-report measure and does not require a great deal of time or researcher/clinician input. Perhaps for this reason, the SAS is more widely used than the LIFE, particularly for cross-sectional research. The SAS includes 54 self-report items and assesses an individual’s ability to function and derive satisfaction from various life roles. Social role areas include work, student and leisure activities, relationships with extended family, role as marital partner, parental role, and role within the family unit (including items related to economic functioning). Within each area, items assess performance, interpersonal conflict, interpersonal relations, and satisfaction. Items are scored on a five-point scale and means are calculated for each relevant role. Higher scores correspond to greater impairment. An overall social functioning score can be obtained by averaging all items answered. Because an individual may not have answered questions for each role (e.g., s/he may
not be a parent), this overall score accounts for functioning in only those domains that are relevant. Mean scores can be converted to T-scores that are normed by gender. The SAS has satisfactory internal consistency, test-retest reliability, and evidence of validity (divergent, discriminant, convergent, external, and concurrent; Weissman et al., 2001).

Global Assessment of Functioning

The Global Assessment of Functioning (GAF; APA, 2000) is used in the CLPS study and also more broadly. The scale is intended to capture overall functioning across social and non-social domains. Individuals with extensive training on using the GAF show strong interrater reliability (e.g., Goldman, Skodol, & Lave, 1992). However, GAF scores have shown poor interrater reliability among raters with less training, including clinical practitioners (Sonesson, Tius, & Arvidsson, 2010). Problematically, experience alone does not improve GAF reliability among clinicians to any great extent (Spengler et al., 2009). The GAF has also shown low clinical utility (Aas, 2010). These are among the problems that led to the deletion of the GAF scale from DSM-5 (APA, 2013). Aside from these concerns, the general scope of the GAF is not specific to social functioning. Low scores on the GAF may thus indicate other forms of distress or impairment and cannot be assumed to capture any particular aspect of social functioning for a given individual. Additionally, the GAF is highly symptom-focused. Highest GAF scores are given to individuals with no symptoms. Lowest GAF scores are given to individuals with symptoms that specifically present danger to self or others (e.g., suicidality, delusions). Thus different types of symptoms are a significant determinate of GAF ratings.

Experience Sampling

Other researchers have sought to measure social functioning without the additional confound of memory bias by collecting data through experience sampling or daily diaries.
Experience sampling methods require an individual to respond to a series of questions when prompted by a digital device at various times throughout the day. This method has good ecological validity, as it gathers information from participants as they are engaged in actual activities of daily life. Daily diary studies require that participants respond to a series of questions or prompts on a daily basis. Although this method reduces the likelihood of memory bias, it is still possible for the 24 hours between any two diary entries to contribute to biased or distorted reports of functioning. This day-long delay may be long enough for recall of social functioning to be distorted by events that have occurred in that interval or by mood at time of reporting. Recent work has used experience sampling to examine the course of borderline symptoms, including mood variability (Houben et al., 2015; Nisenbaum et al., 2010) and impulsivity (Chapman, Rosenthal, & Leung, 2009). With regard to social functioning, experience sampling and diary methods have been used to examine interpersonal triggers (Miskewicz et al., 2015) and relationship satisfaction (Kuhlken et al., 2014) in BPD. No work to date has used these methods to investigate overall social functioning in BPD. This may be due to the high resource burden of these types of assessments.

**Other Measures**

Other studies focus on social functioning within one defined role or domain. Romantic relationship dysfunction, for example, has received a great deal of attention (e.g., Daley, Burge, & Hammen, 2000). Finally, some studies use the interpersonal instability criterion of BPD as their measure of social function (e.g., Zanarini et al., 2003). Presumably those individuals who endorse interpersonal instability are likely to have low social functioning in some domains. However, it is possible at least in principle to endorse this criterion and others while maintaining role functioning in other important domains. Additionally, using this criterion as a marker of
social functioning implicitly assumes that individuals who have BPD but who not endorse this criterion are functioning well.

**Expanding the Scope of Social Functioning**

The use of disparate measures and enduring problems with our current measurement of social functioning may be due in part to a lack of conceptual consensus. The terms “interpersonal problems” or “interpersonal dysfunction” are used to refer to a wide range of specific interpersonal processes. One study may refer only to interpersonal hostility (e.g., Minzenberg et al., 2006) while another may use one of these terms to refer to rather disparate constructs. For example, some have examined the association of borderline features with specific interpersonal outcomes, including marriage likelihood (e.g., Swartz, Blazer, George, & Winfield, 1990), divorce or other dissolution (e.g., Bouchard, Sabourin, Lusier, & Villenvue, 2009), or with specific aspects of interpersonal interactions, including relationship conflict (e.g., Chen et al., 2004) and satisfaction (e.g., Bouchard et al., 2009). Other processes of interest in the study of BPD could also be considered aspects of social functioning, including rejection sensitivity, social cognition, and attachment.

The current literature on social functioning in BPD was reviewed and gaps in measurement were identified. Currently, two measures offer excellent coverage of social functioning across domains: the LIFE and the SAS. The study of social functioning or interpersonal dysfunction in BPD must necessarily be broad in scope; inclusion of multiple domains of social functioning is likely important for our understanding of the distress and impairment that characterize the disorder. All of the current measures be they self-report, interview, or experience sampling, are necessarily limited in scope, and often do not measure all aspects of social functioning that may be of interest in the study of BPD. The LIFE and the SAS
are not immune to this criticism. Here we sought to broaden the scope defined by current assessment tools to include several additional aspects of social functioning.

**New Measure Development**

To this end, I developed and pilot-tested measures intended to capture aspects of social functioning that are often neglected by current assessment tools. No current instruments offer adequate coverage of prosocial and antisocial behaviors, the frequency at which individuals initiate positive or negative interactions, or the frequency at which individuals feel that others initiate positive or negative social interactions with them.

Common conceptions of BPD often include the notion that individuals with the disorder are manipulative. Whether this is an appropriate conceptualization or an inappropriately stigmatizing one is beyond the scope of this study. However, there is evidence that BPD shares important overlaps with antisocial personality disorder (APD), a disorder that is well characterized by manipulation and other social behaviors that violate societal expectations. BPD and APD share a number of features, including impulsivity, negative affect, and aggression. They also share mutual genetic (Kendler et al., 2008; Torgersen et al., 2008) and environmental risk factors (e.g., childhood trauma) and common comorbidities (e.g., substance use disorders). The disorders commonly co-occur both in the general population (Trull et al., 2013) and clinical samples (Becker et al., 2000). Men with BPD may be incorrectly diagnosed with APD based on gender biases and the common occurrence of antisocial traits in men who have BPD (Sansone & Sansone, 2011). Some have argued that this contributes to an over-diagnosis of APD in men (Ford & Widiger, 1989). Conversely, elevated psychopathy is common in BPD, and some have argued that BPD may be a phenotypic expression of psychopathy that may present as BPD primarily in women (Sprague et al., 2012). In other words, psychopathy may manifest differently
depending on gender. If BPD and APD share vulnerabilities and similar features, one may reasonably expect that individuals with borderline traits may perform more antisocial behaviors. Frequency of antisocial behavior may indicate one aspect of social functioning. Here I sought to develop a measure of mild to moderate antisocial behaviors as a marker of social functioning.

In contrast to antisocial behaviors, prosocial behaviors are an important aspect of human social exchange. Prosocial behaviors include acts that benefit other individuals, groups of individuals, or society more broadly. These acts are performed voluntarily, even when there is no clear benefit for the actor. Frequency of prosocial behaviors may be regarded as one marker of social functioning. Additionally, prosocial behaviors may be motivated in part by empathy. Impairments in empathy have been at least inconsistently linked to BPD (Jeung & Herpertz, 2014), and thus one might reasonably expect that borderline features are associated with lower frequency of prosocial acts. Perhaps because of the warranted focus on negative affectivity and negative aspects of social functioning in BPD, no work to date has examined this association. Here I sought to measure a range of prosocial acts and to determine how borderline traits are associated with overall prosociality.

Finally, the majority of social functioning measures that are currently used address role functioning (e.g., the SFQ includes an item that reads “I get on well with my family and other relatives”) or broad social processes (e.g., the IIP-PD-25 includes an item that reads “I am too envious of other people”) rather than specific, concrete behaviors. When a participant rates how well they get on with their family or other relatives, it is unclear how often that participant initiates contact with those relatives. It is also unclear how often those relatives initiate contact with the participant. Is this a participant who gets on well with relatives but who is generally isolated and see them only on occasion? Or is this a participant who consistently initiates
contact? Similarly, a participant may feel that she is too envious of other people, but this may or may not act as a barrier for that participant to initiate interactions with others. Enviousness may also manifest in interpersonal behaviors that lead others to initiate contact less often, or it may not. The overall message here is that measuring the actual frequency of social contact of various forms may provide additional information about social functioning that is not currently captured by validated assessment tools. Here I developed a scale designed to measure the frequency of interactions initiated by the participant and toward the participant by others.

The current goal was to test whether measures of prosocial and antisocial acts, as well as measures of interactions initiated by and toward participants, were reliably related to borderline personality features. The current chapter includes initial pilot testing of new measures to gain a better understanding of how they relate to borderline features, and with an eye toward using these measures in Chapter 6 to examine how trustworthiness appraisals processes relate to social functioning. Here I measured both borderline and schizotypal personality traits in the same individuals. Schizotypal traits were included for the sake of comparison. Although this initial study was largely an exploratory pilot test of how these new measures relate to borderline pathology, I made several predictions. I predicted that both borderline and schizotypal traits would be associated with lower frequency of positive interaction initiation by participants, but that only borderline personality would be associated with higher frequency of negative interaction initiation. I predicted that both sets of traits would predict lower reports of positive social interaction initiations from other people and higher reports of negative social interaction initiations from other people. I also predicted that both sets of traits would negatively predict satisfaction with social interactions. Finally, I predicted that borderline personality traits would
be uniquely related to antisocial behaviors, and that neither borderline nor schizotypal traits would be associated with prosocial behaviors.

**Method**

Three scales were created through rational item selection. Potential items were offered by a total of eight individuals with research training in scale development, clinical experience with borderline personality, or both. The final three scales are provisionally entitled the *Nobody’s Perfect Questionnaire* (NPQ), the *Being Nice Questionnaire* (BNQ) and the *Social Interaction Initiation Scale* (SIIS).

In an effort to determine how demand characteristics might influence responses on these measures and to eliminate this influence, I included items intended to measure honesty in self-report, and no participants were required to respond to any questions they did not wish to answer. Additionally, I included an item to capture general satisfaction with social leisure time, with the expectation that those individuals who initiate positive interactions and experience more positive initiations from others would have higher global satisfaction with their social leisure time. Full scales are included in Appendices C-E.

**Participants**

Participants were recruited using Amazon’s Mechanical Turk. The study title was listed as “15-30 Minute Social Interactions Survey” and participants were told that the “study aims to better understand social experiences that many of us have on a regular basis as well as behaviors that some of us do.” They were also instructed that

By clicking the link below, you will be taken to a survey that runs through a program called Qualtrics, and asks you a series of questions about yourself. These questions will ask about social experiences that you may or may not have on a regular basis. Some questions will ask about good or bad things that you may have done in the past. We recognize that no one is “all good” or “all bad” and encourage your honesty on these
questions. We ask that you complete the questions in one sitting, if possible. They should take about 15 minutes.

Criteria for inclusion were that participants spoke English as their first language and were at least 18-years-old.

Mechanical Turk (MTurk) is an online crowd-sourcing platform that allows registered online workers to complete computerized tasks and surveys in return for small compensation. “Requesters” looking to hire MTurk “workers” can post descriptions of tasks, including demands and eligibility requirements. Tasks are referred to as Human Intelligence Tasks (HITs), and are displayed to registered MTurk workers, who can choose to accept and complete any posted task for which they qualify. MTurk is efficient and low-cost and provides researchers with access to a large participant pool. MTurk has approximately 500,000 registered workers across a 190 countries. A majority of this population is based out of the United States and India, with approximately 47% of the workforce located in the USA, 34% located in India and the remaining 19% distributed across the rest of the world (Paolacci et al., 2010; Paolacci & Chandler, 2014). The current study was made available only to registered workers within the United States.

MTurk participants are more demographically diverse than college samples, with a greater representation of non-White participants (36%), and a marginally greater representation of females (55%; Paolacci & Chandler, 2014; Buhrmester et al., 2011; Gosling et al., 2004). MTurk workers and general community samples differ in ways congruent to differences typically observed between frequent Internet users and non-internet users. MTurk workers are less extraverted, more socially anxious, and less emotionally stable (Goodman, Cryder, & Cheema, 2013; Shapiro et al., 2013; Goodman et al., 2013). However, other work has found that MTurk workers do not differ from community populations in level of emotional dysregulation (Shapiro et al., 2013).
Evidence suggests that MTurk can be used as a reliable and valid platform for survey studies, studies including various behavioral tasks (e.g., the Stroop; Crump et al., 2013; Buhrmester et al., 2011). Studies on the reliability and validity of data collected via MTurk are promising, with findings of test-retest reliability ranging from $r = 0.80$ to $r = 0.94$ (Buhrmester et al., 2011), and internal consistencies of validated questionnaires ranging from $\alpha = .73$ to $\alpha = .93$. Other work has shown that data collected via MTurk yields similar or greater internal consistencies than data collected using populations of college students (Behrend et al., 2011). Within the realm of clinical populations, MTurk data have yielded similarly strong indices of reliability and validity, including on the Beck Depression Inventory ($\alpha = 0.90$, $r = 0.87$), Beck Anxiety Inventory ($\alpha = 0.93$), and Liebowitz Anxiety Scale ($\alpha = 0.97$; Behrend et al., 2011). In the same study, analysis of clinical symptom reporting showed that the levels of clinical depression (approximately 5%) and generalized anxiety (approximately 2.9%) reported by the MTurk sample were comparable to national U.S. prevalence rates (7% for depression and 3.1% for anxiety; Kessler, Chiu, Demler, & Walters, 2005). MTurk users reported significantly higher rates of social anxiety (50.5%) than the general population (6.8%). This difference accords with previous studies that report higher levels of social anxiety in individuals with greater Internet usage (Bargh & McKenna, 2004; Cole & Hooley, 2013; Shepherd & Edelmann, 2005; Stevens & Morris, 2007; Weidman et al., 2012).

Participants were 100 individuals (60% female). Each participant was paid $0.50 for their time and was entered in a raffle to win one of 5 $10 gift cards. Participant age ranged from 19-75 years ($M = 40.8$, $SD = 15.6$). Participants were married or living with a partner (48%), divorced/separated (11%), single (40%), or widowed (1%). All participants endorsed high school
education or equivalent, 40% had attended some college, and 30% held Bachelors’ degrees. A total of 23% indicated that they were currently unemployed.

**Measures**

**New Social Functioning Measures.** As described above, the NPQ, the BNQ, and the SIIS were included. Full measures are included in Appendices C-E.

**SNAP-2.** The Schedule for Nonadaptive and Adaptive Personality-2 (SNAP-2; Clark, 2006) is described in more detail in Chapter 2. Participants completed the items from the diagnostic scales designed to measure borderline and schizotypal personality. The schizotypal scale includes a subset of 25 items that assess 8 of the 9 criteria for schizotypal personality disorder (one criterion, disorganized speech or though, requires direct observation rather than self-report). The borderline scale includes a subset of 33 items that assess all 9 of the criteria for borderline personality disorder. The two scales share 4 items in common.

**Results**

**Endorsement of Borderline and Schizotypal Personality**

Borderline and schizotypal personality were endorsed at high rates. Based on diagnostic criteria as derived from the SNAP-2, 6% of the current sample met diagnostic threshold for BPD (5 or more criteria). Another 42% of the sample endorsed 1-4 borderline criteria, leaving 52% of the sample endorsing no criteria. With regard to schizotypal criteria, 4% of the sample met diagnostic criteria (5 or more criteria). Notably the SNAP-2 is unable to assess one criterion for SPD, as observation is necessary to determine whether a participant presents with odd thinking or speech. Because participants are essentially missing one opportunity to meet a criterion, the prevalence of SPD in this sample may actually be higher. Another 53% of the sample endorsed 1-4 criteria, leaving 43% of the sample endorsing no schizotypal personality criteria. The
The number of criteria endorsed is outlined in Tables 5.1 and 5.2. The frequencies for specific symptoms endorsed are included in Tables 5.3 and 5.4.

**Table 5.1. BPD Prevalence: 6%.**

<table>
<thead>
<tr>
<th>Number of Diagnostic Criteria Met</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 5.2. SPD Prevalence: 4%.**

<table>
<thead>
<tr>
<th>Number of Diagnostic Criteria Met</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>43</td>
</tr>
<tr>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

*Note.* Criterion 5 (odd thinking or speech) is observational and was not assessed, so participants cannot endorse all 9 criteria.
Table 5.3. Specific Borderline Symptoms.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Symptom</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fears of abandonment</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Interpersonal instability: idealizing and devaluing</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Identity disturbance</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>Impulsivity</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>Self-harm</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Affective instability</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>Chronic feelings of emptiness</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>Inappropriate or intense anger</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Stress-related paranoia or dissociation</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 5.4. Specific Schizotypal Symptoms.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Symptom</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ideas of reference</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Odd beliefs or magical thinking</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Unusual perceptual experiences</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>Odd thinking and speech</td>
<td>Observational</td>
</tr>
<tr>
<td>5</td>
<td>Suspiciousness or paranoid ideation</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Inappropriate or constricted affect</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>Odd or eccentric behavior or appearance</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>Lack of close friends or confidants</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>Excessive social anxiety</td>
<td>10</td>
</tr>
</tbody>
</table>

**Internal Consistency**

Internal consistency was analyzed using Cronbach’s Alpha. Internal consistencies were strong for the SIIS, including subscales measuring the initiation of positive and negative interactions by the participant and by others toward the participant. Subscales include positive interactions initiated by the participant ($\alpha = .798$), negative participant-initiated interactions ($\alpha = .934$), other-initiated positive interactions ($\alpha = .770$), and other-initiated negative interactions ($\alpha$...
Finally, the BNQ and the NPQ also achieved high internal consistency ($\alpha = .899$ and $\alpha = .917$), respectively. No scales could be improved by item deletion.

**Age and Social Functioning**

Older age was negatively associated with borderline criteria $r(98) = -0.34$, $p < .01$, while age was unrelated to the number of schizotypal criteria endorsed $r(98) = -0.19$, $p > .05$. Older age was further associated with performance of prosocial behaviors $r(98) = 0.30$, $p < .01$ even when controlling for borderline criteria, $\beta = 0.323$, $p = 0.002$. Age was unrelated to performance of antisocial behaviors $r(98) = -0.15$, $p > .05$. Age was not associated with general social satisfaction $r(98) = -0.01$, $p > .05$, the frequency with which individuals initiated positive interactions with others $r(98) = -0.11$, $p > .05$ or with which others initiated positive interactions with participants $r(98) = -0.15$, $p > .05$. However, age was negatively related to the initiation of negative interactions by the participant $r(98) = -0.37$, $p < .01$ even when controlling for borderline criteria, $\beta = -0.242$, $p = 0.012$. Age was also negatively related to the initiation of negative interactions from others $r(98) = -0.36$, $p < .01$, but when controlling for borderline criteria this association is only borderline significant, $\beta = -0.169$, $p = 0.052$.

**Associations with Borderline Criteria**

Borderline and schizotypal criteria were strongly correlated, $r(98) = 0.75$, $p < .01$. Borderline criteria were also significantly correlated with performance of “bad” behaviors on the NPQ, such that those who endorsed more borderline traits also endorsed higher frequency bad behavior, $r(98) = 0.43$, $p < .01$. This was true even when accounting for schizotypal traits and age, ($\beta = 0.826$, $p < 0.001$). Borderline criteria were not associated with the frequency with which participants engaged in prosocial behaviors, $r(98) = -0.07$, $p > .05$. As expected, borderline criteria were also associated with general social satisfaction, such that those who
endorsed more criteria were less satisfied with their free time, \( r(98) = -0.24, p < .05 \). Notably, this relationship was not significant when controlling for the contribution of schizotypal criteria to social satisfaction, \( \beta = -0.94, p = 0.53 \). When age was entered as a covariate without schizotypal criteria, borderline criteria remained significantly and negatively associated with general social satisfaction, \( \beta = -0.284, p = 0.008 \). Endorsement of borderline criteria was also associated with more frequent initiation of negative social interactions, \( r(98) = 0.44, p < .01 \) (even when controlling for schizotypal criteria and age, \( \beta = 0.572, p < 0.001 \)), and more frequent experiences of negative interactions initiated by others \( r(98) = 0.59, p < .01 \) (even when controlling for schizotypal criteria and age, \( \beta = 0.592, p < 0.001 \)). Borderline criteria were not correlated with the frequency with which participants initiated positive social interactions \( r(98) = 0.02, p > .05 \), or the frequency with which they experienced positive social interactions initiated by others \( r(98) = 0.09, p > .05 \). Table 5.5 presents zero-order intercorrelations of personality features and measures of social functioning.

**Item-Level Associations with Borderline Criteria**

On the item level, borderline criteria were associated with a single prosocial behavior on the BNQ: “taking the blame for someone else to help them” (\( r(98) = 0.26, p = .009 \)). This association remained significant controlling for age and schizotypal criteria, \( \beta = 0.505, p = 0.002 \). This is a particularly subjugating prosocial behavior, and it would be helpful to determine whether this association remains significant when controlling for features of dependent personality disorder, which was not included in the current study. In terms of antisocial behaviors on the NPQ, borderline criteria were significantly and positively associated with more than half of the scale items (items 3, 6, 8, 12, 13, 14, 15, 17, 18, 20, 21, 24, 25, 26, and 28). The strongest correlations (above .40) were with items that read “cheating on a test” (\( r(98) = 0.43 p <
.001), “passing by someone who needs help that I can give” ($r(98) = 0.41, p < .001$) and “stealing something from a person I know” ($r(98) = 0.44, p < .001$). These associations remained significant when controlling for age and schizotypal criteria.

With regard to initiation of positive interactions, borderline criteria were positively associated with “initiated plans with friends such as dinner, going out, or attending a party” ($r(98) = 0.23, p = .02$), “made a new friend” ($r(98) = 0.27, p = .006$), and “flirted with someone, or asked someone out on a date” ($r(98) = 0.23, p = .02$). Notably, the last two of these three items involve approach behavior in the context of unknown others. Borderline criteria were negatively associated with the frequency at which individuals “spoke with a friend in person” ($r(98) = -0.21, p = .039$). The positive associations survived controls for schizotypal criteria and age. However, the association of borderline criteria with frequency of speaking to friends in person was nonsignificant when age and schizotypal criteria were controlled, $\beta = 0.258, p = 0.082$.

Borderline criteria were significantly and positively associated with the initiation of negative interaction behaviors. The one item not associated with borderline criteria read “wished that you were spending time alone” ($r(98) = 0.10, p = .32$). This is consistent with the yearning for social connection that is often observed clinically in the disorder, as well as the common report that chronic feelings of emptiness are more difficult when patients are alone.

With regard to the experience of positive interaction initiations from others, borderline criteria were positively associated with “been invited to a planned activity with a group” ($r(98) = -0.31, p = .002$) and “been asked on a date, or was flirted with” ($r(98) = 0.26, p = .01$). Borderline criteria were negatively associated with “been called by a family member” ($r(98) = -0.21, p = .037$). Positive associations remained when controlling for age and schizotypal criteria.
However, these controls made the association between borderline criteria and “been called by a family member” nonsignificant, $\beta = 0.07, p = 0.66$.

Borderline criteria were positively associated with every item on the other-initiated negative behavior subscale of the SIIS. In other words, individuals with more borderline symptoms endorsed experiencing more frequent negative social overtures from others across all items. The strongest association was with “felt someone you did not know was being rude to you” ($r(98) = 0.60, p < .001$). With one exception, all items remained significantly associated with borderline criteria when age and schizotypal criteria were controlled. “Not been invited to a social gathering which your friends had organized” was not significantly associated with borderline criteria in this case, $\beta = 0.27, p = 0.071$.

**Associations with Schizotypal Criteria**

As expected, in contrast to borderline criteria, schizotypal criteria were not significantly correlated with performance of “bad” behaviors on the NPQ, $r(98) = 0.11, p > .05$, nor were they associated with the performance of prosocial behaviors, $r(98) = -0.13, p > .05$. Also as expected, schizotypal criteria were associated with global social satisfaction, such that those who endorsed more criteria were less satisfied with their free time, $r(98) = -0.27, p < .01$. However, this relationship was not significant when controlling for the contribution of borderline criteria to social satisfaction, $\beta = -0.198, p = 0.19$. The correlation of global social satisfaction with total number of schizotypal and borderline criteria endorsed was significant, $r(98) = -0.27, p = .006$. Endorsement of schizotypal criteria was also associated both with less frequent participant-initiated $r(98) = -0.25, p < .05$ and other-initiated, $r(98) = -0.21, p < .05$ positive social interactions. Both of these associations remained significant when controlling for borderline criteria ($\beta = -0.608, p < 0.001$ and $\beta = -0.637, p < 0.001$, respectively). Like with borderline
criteria, schizotypal criteria were also associated with more frequent experiences of negative interactions initiated by others \( r(98) = 0.40, p < .01 \). However, this effect was not significant when controlling for borderline criteria, \( \beta = -0.109, p = 0.38 \). Schizotypal criteria endorsed were not correlated with the frequency with which participants initiated negative social interactions \( r(98) = 0.19, p > .05 \). Table 5.5 presents zero-order intercorrelations of personality features and measures of social functioning.

**Item-Level Associations with Schizotypal Criteria**

On the item level, schizotypal criteria were negatively associated with two prosocial behaviors on the BNQ: “donating money to charity” \( r(98) = -0.20, \ p = .045 \) and “calling someone on the phone to see how they are doing” \( r(98) = -0.31, \ p = .002 \). The association with “donating money to charity” was no longer significant controlling for borderline criteria \( \beta = -0.172, p = 0.258 \). However, the association with “calling someone on the phone to see how they are doing” remained significant \( \beta = 0.102, p = 0.01 \). In terms of antisocial behaviors on the NPQ, schizotypal criteria were significantly and positively associated with four items: “calling in to work sick when I’m not sick” \( r(98) = 0.24, p = .015 \), “saying something mean to hurt someone else” \( r(98) = 0.25, p = .011 \), “passing by someone who needs help I can give” \( r(98) = 0.22, p = .029 \), and “purposely avoiding someone so I don’t have to talk to them” \( r(98) = 0.30, p = .003 \). None of these associations remained significant when controlling for borderline criteria. Schizotypal criteria were negatively associated with “littering” \( r(98) = -0.24, p = .018 \), and this remained significant when controlling for borderline criteria \( \beta = -0.438, p = 0.04 \).

With regard to initiation of positive interactions, schizotypal criteria were not positively associated with any items. However, they were negatively associated with “spoken to a friend via phone or email” \( r(98) = -0.37, p < .001 \), and “spoken with a friend in person” \( r(98) = -0.40, p \)
< .001). This may be reflective of a smaller pool of friendships, a characteristic of schizotypal personality disorder. Both associations survive controls for borderline criteria.

Schizotypal criteria were significantly and positively associated with the initiation of 5 negative interaction behaviors: “had an argument with someone you are close to” ($r(98) = 0.24, p = .02$), “were unsatisfied with the quality of your social interactions” ($r(98) = 0.22, p = .03$), “walked away from a social interaction feeling bad or awkward about it” ($r(98) = 0.20, p = .04$), “regretted the way you behaved in a social situation” ($r(98) = 0.22, p = .028$), and “criticized someone unnecessarily” ($r(98) = 0.24, p = .02$). None of these interactions remained significant when controlling for borderline symptoms.

With regard to the experience of positive interaction initiations from others, schizotypal criteria were negatively associated with “been called by a family member” ($r(98) = -0.28, p = .004$) and “received socially related emails or text messages” ($r(98) = -0.29, p = .004$). The former did not survive controls for borderline criteria ($\beta = -0.29, p = 0.053$), but the latter remained significant ($\beta = -0.442, p = 0.003$).

Schizotypal criteria were positively associated with 7/10 items item on the other-initiated negative behavior subscale of the SIIS. In other words, individuals with more schizotypal symptoms endorsed experiencing more frequent negative social overtures from others across 70% of items. Significant correlations ranged from $r = 0.22$ to $r = 0.46$. Only one item remained significantly associated with schizotypal symptoms when controlling for borderline symptoms: “had someone cancel plans you initiated” ($\beta = -0.39, p = 0.006$). Notably, the direction of the association changed once borderline symptoms were entered into the model: schizotypal symptoms, in the absence of borderline symptoms, are associated with less frequent experiences of others cancelling plans.
Scale Intercorrelations

Frequency of prosocial and antisocial behaviors were positively correlated—those who reported initiating more prosocial behaviors also reported performing more negative, antisocial behaviors $r(98) = 0.33, p < .01$. Neither prosocial $r(98) = 0.03, p > .05$ nor antisocial behaviors $r(98) = -0.00, p > .05$ were associated with general social satisfaction, but both were associated with the frequency at which participants endorsed initiating positive interactions with others $r(98) = 0.34, p < .01$ and $r(98) = 0.38, p < .01$, which in turn was associated with higher general social satisfaction $r(98) = 0.33, p < .01$. Prosocial behaviors were not associated with participant-initiated negative social interactions $r(98) = -0.05, p > .05$, but participant-initiated negative social interactions were associated with more frequent performance of antisocial behaviors ($r(98) = 0.56, p < .01$). These negative social interactions were not associated with general social satisfaction $r(98) = -0.08, p > .05$. Other-initiated positive social interactions were positively associated with participant prosocial $r(98) = 0.37, p < .01$ and antisocial $r(98) = 0.41, p < .01$ behaviors, leisure satisfaction $r(98) = 0.25, p < .01$, and both participant-initiated positive $r(98) = 0.83, p < .01$ and negative social interactions $r(98) = 0.40, p < .01$. Other-initiated negative social interactions were not associated with participant prosocial behavior $r(98) = -0.06, p > .05$, were positively associated with antisocial behaviors $r(98) = 0.47, p < .01$, and participant-initiated positive $r(98) = 0.24, p < .05$ and negative behaviors $r(98) = 0.82, p < .01$. Other-initiated negative interactions were negatively related to general social satisfaction $r(98) = -0.26, p < .05$. 

121
**Discussion**

Overall, these preliminary results indicate that the assessment’s subscales have strong internal consistency and perform in predicted ways in relation to borderline and schizotypal traits. As expected, borderline features were associated with performance of antisocial behaviors, while schizotypal features were not. Neither set of personality features were associated with prosocial acts. In terms of positive and negative self- and other-initiated interactions, borderline
features were associated with higher rates of self-initiated negative interactions, as might be expected given the clinical presentation of the disorder. Borderline features were also associated with the experience of more other-initiated negative behaviors. The current data are insufficient to determine whether a negativity bias might influence this report. However, this is largely to be expected given that families and friends of individuals with borderline traits often report having difficulty interacting positively with them, and the feeling of needing to “walk on eggshells.” Additionally, one might expect a pattern of reciprocal negativity that perpetuates both self- and other-initiated negative behaviors.

Schizotypal features were not related to self-initiated negative interactions. This is congruent with the clinical manifestation of schizotypal personality disorder, which is characterized by excessive social anxiety, suspiciousness, and eccentricity, rather than the negative interpersonal style characteristic of borderline personality disorder. The differentiation between schizotypal features and borderline features on the self-initiated positive interaction subscale is also to be expected given high levels of social anxiety in schizotypal personality. Schizotypal features were also associated with a lower frequency of other-initiated positive behaviors. While not surprising, this effect is somewhat difficult to account for. It may be the case that others have difficulty knowing how to relate positively to individuals who display eccentricity or paranoia. It could also be the case that this captures a level of social withdrawal that differentiates borderline and schizotypal personality. Both sets of personality features were associated with lower global social satisfaction. However, when controlling for schizotypal criteria, the association of borderline criteria and social satisfaction was no longer significant. The reverse is also true: when controlling for borderline criteria, the association of schizotypal criteria and social satisfaction was nonsignificant. The association of total personality disorder
criteria (schizotypal and borderline criteria combined) was significantly associated with general social satisfaction. This suggests that global social satisfaction is best predicted by personality disorder broadly rather than specific personality disorders themselves.

Overall, there is good differentiation of how the scales perform with regard to borderline versus schizotypal personality features. As expected, borderline features are associated with more of the antisocial or negative aspects of social behaviors and are not correlated with any of the positive aspects of social behavior. On the other hand, schizotypal features are not broadly associated with antisocial or negative social behaviors. These scales thus appear to capture with specificity the hostility and negative style that characterizes borderline personality.

Results for age are also largely as expected. Age was related to borderline features but not schizotypal features, such that older individuals reported fewer borderline features. This is consistent with the notion that although genetic vulnerabilities to borderline features may remain unchanged across the lifespan, environmental vulnerabilities may decrease with age, and particularly in the context of adult relationships (Reichborn-Kjennerud et al., 2015). In comparison to BPD, schizotypal personality disorder has a larger genetic contribution, and thus may be more stable over time (Kendler et al., 2015). Age was not associated with global social satisfaction or self-initiation of positive interactions. The lack of association of age and self-initiation of positive interactions indicates that the interactions selected are relevant to individuals across the lifespan, and not activities that older or younger people engage in at different rates. Age was associated with self-initiation of negative interactions, such that older individuals reported initiating fewer negative interactions. This was true even controlling for borderline symptoms. Older individuals also reported fewer other-initiated negative interactions, but this effect was nonsignificant when controlling for the effect of borderline symptoms. This
suggests that borderline symptoms drive other-initiated negative interactions, and that these interactions may become less frequent with age but likely as a result of decreases in borderline symptomatology rather than something more inherent to the aging process.

Broadly, results validate that these scales are potentially useful in the study of borderline personality disorder. Better understanding what moderates or mediates the association of borderline personality with antisocial acts and negative self- and other-initiated interactions may prove fruitful for understanding and changing the interpersonal dysfunction that characterizes the disorder. Additionally, these measures serve to broaden the scope of social functioning as currently conceptualized. Although the current data are insufficient to fully support the idea that these scales may be less influenced by negativity bias, it is possible that clearly defined acts/behaviors may be less open to distortion than scales that ask broader questions that are more open to interpretive differences. Future work would do well to address the question of what kinds of social functioning measures are least contaminated by negativity bias. What is clear from the current study is that borderline features are associated with more antisocial and negative social behaviors. Whether these types of behaviors are related to trust processing biases is investigated in the next chapter.
Chapter 6: Trustworthiness Appraisal and Social Functioning
Introduction

Trust processing biases for individuals with borderline features have been identified in past literature and in the preceding chapters. There is now robust evidence that compared to individuals without borderline features, individuals with borderline features rate others as less trustworthy. In Chapter 3, we outlined evidence that negative affective information exerts exaggerated influence on trustworthiness appraisals made by individuals with borderline features. In Chapter 4, we showed that the perception of a face as untrustworthy may contribute to incorrect memory judgments about that face, specifically for individuals with borderline features. This information is important to our understanding of borderline pathology and the influence of affective information on social cognitive processes. However, this information does not prima facie provide any indication of whether these laboratory-observed processes contribute to how individuals are actually functioning in their social environments. Currently we know of one ongoing study by Miano and colleagues that is considering the role of trust appraisals in romantic relationship functioning. However, no published studies thus far have attempted to relate trust processing biases found in the laboratory directly to real-world social functioning.

Important obstacles exist in relating trust processing biases to real-world social function. In Chapter 5 we outlined some of these inherent difficulties and piloted new measures of social functioning. Here, in Chapter 6, we used these measures in addition to previously validated measures to examine whether trust appraisal processes observed in Chapter 3 related to self-reported social functioning. Although this chapter is largely exploratory in nature, we did expect that trust appraisal bias and influence of negative affective information on trust appraisals would be broadly predictive of reported social functioning. More specifically, we predicted that (1) the borderline group would report lower social-functioning across all measures. Similarly, we expected that borderline features would be associated with lower social functioning across all
measures; (2) a tendency to rate others as less trustworthy would be related to poorer social functioning for the sample overall; (3) trustworthiness ratings would mediate the relationship between borderline features and social functioning; (4) the magnitude of negative affective priming influence on trustworthiness appraisals would predict poorer social functioning for the sample overall; and (5) negative affective priming influence would mediate the relationship between borderline features and social functioning.

**Method**

The same participants who completed the tasks described in Chapter 3 and Chapter 4 also completed a series of measures designed to measure social functioning. Sample characteristics for these participants are described in detail in Chapter 2. Full descriptions of measures, including the social functioning measures used here, are included in Chapter 2 and Chapter 5. Social functioning measures are listed in Table 6.1.

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Citation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nobody's Perfect Questionnaire (NPQ)</td>
<td>New: See Appendix</td>
<td>Common antisocial behaviors</td>
</tr>
<tr>
<td>Being Nice Questionnaire (BNQ)</td>
<td>New: See Appendix</td>
<td>Common prosocial behaviors</td>
</tr>
<tr>
<td>Social Interaction Initiation Scale (SIIS)</td>
<td>New: See Appendix</td>
<td>Initiation of positive and negative interactions with others, either by the participant or by another person</td>
</tr>
<tr>
<td>General Leisure Satisfaction</td>
<td>New: See Appendix</td>
<td>Single-item report of general satisfaction with leisure time</td>
</tr>
<tr>
<td>Social Function Questionnaire (SFQ)</td>
<td>Tyrer et al. (2005)</td>
<td>Brief, 8-item measure</td>
</tr>
<tr>
<td>Experiences in Close Relationships Scale (ECR)</td>
<td>Fraley et al. (2000)</td>
<td>Scale was modified to refer to relationships generally rather than romantic relationships specifically. Scale altered from 7-point likert to 5-point likert. Included in Appendix.</td>
</tr>
</tbody>
</table>

*Note.* See Chapters 2 and 5 for further details.
Results

Prosocial and Antisocial Behaviors

Correlations among social functioning measures are presented in Table 6.2. The commission of prosocial acts was measured using the Being Nice Questionnaire (BNQ). The BNQ is described in greater detail in Chapter 5 and is included in Appendix E. This measure was unrelated to any other measure of social functioning. This suggests that individuals can be doing poorly in multiple domains of social functioning but still engage in prosocial behavior. It may also be the case that very serious pathology is required for deficits in this area of social functioning—the current sample may not be severe enough to capture changes in prosocial behavior. Alternatively, the lack of association between prosocial acts and other measures of social functioning, particularly contrasted with the association of bad acts and other measures described below, suggests that this measure may not be a relevant indicator of social functioning in this sample or even generally.

In contrast, the commission of mild antisocial acts measured using was the Nobody’s Perfect Questionnaire (NPQ) was broadly related to poor social functioning across an array of measures. The NPQ is described in greater detail in Chapter 5 and included in Appendix D. Commission of antisocial acts was not related to participant’s report of initiating positive interactions with others on the SIIS, $r(75) = 0.13, p > .05$. This indicates that individuals could endorse performance of both mild antisocial acts and positive engagement with others. The SIIS is described in greater detail in Chapter 5 and included in Appendix C. Performance of mild antisocial acts, was, however, associated with participant initiation of negative interactions ($r(75) = 0.57, p < .01$), initiation of positive ($r(75) = 0.24, p < .01$) and negative ($r(75) = 0.64, p < .01$) social interactions by others toward the participant, self-reports of social satisfaction ($r(75) = 0.31, p < .01$), overall poorer social functioning on the SFQ ($r(75) = 0.40, p < .01$), lower overall
social functioning on the SAS ($r(67) = 0.37$, $p < .01$), and poorer experiences in close
relationships ($r(75) = 0.33$, $p < .01$). The finding that commission of mild antisocial acts is
associated with some positive outcomes (e.g., other-initiated positive interactions), may be the
product of greater overall social engagement. In other words, it may be that individuals who are
most frequently engaged in social exchange have greater opportunity to both commit “bad”
actions and to be the recipient of positive social overtures from others.

**Initiation of Social Interactions**

**Positive interactions initiated by the participant.** The initiation of positive social
interactions by the participant toward others was associated with its counterpart: the initiation of
positive social interactions by others toward the participant, $r(75) = 0.60$, $p < .01$. This is
suggestive of a pattern of positive reciprocal social interaction—those who engage positively
more often are met with more positive engagement in return. The reverse could also be true—
those who experience more positive engagement from others may be more likely to return with
positive social engagement. As might be expected, participants who reported that they initiated
more positive social interactions tended to feel more satisfied with their social lives overall, $r(75)
= -0.34$, $p < .01$. Note that higher self-reported global social satisfaction scores indicate lower
satisfaction, in keeping with the general format of social functioning measures. Finally,
individuals who initiate more positive interactions with others are less likely to experience
attachment concerns as measured with the ECR, $r(75) = -0.41$, $p < .01$. Positive interaction
initiation by participants was unrelated to negative interaction initiation ($r(75) = 0.11$, $p > .05$),
initiation of negative interactions by others toward the participant, ($r(75) = -0.12$, $p > .05$),
overall social functioning on the SFQ ($r(75) = 0.17$, $p > .05$), and overall social functioning on
the SAS ($r(67) = 0.12$, $p > .05$). This pattern of results suggests that positive interaction initiation
by participants may contribute to improved global social satisfaction and more positive social exchange. However, it is largely unrelated to overall social functioning on measures that capture more than individual exchange and general satisfaction. In other words, initiation of positive interactions may be unrelated to how well an individual is able to function at home, at school, at work, or other contexts.

**Negative interactions initiated by the participant.** In contrast, participant initiation of negative social interactions was broadly predictive of poor social functioning across an array of measures. Possibly due to social reciprocity effects, it was associated with more frequent negative interactions initiated by others toward the participant, $r(75) = 0.77, p < .01$. It was also associated with lower leisure satisfaction ($r(75) = 0.47, p < .01$), lower overall social functioning on the SFQ ($r(75) = 0.64, p < .01$) and the SAS ($r(67) = 0.56, p < .01$), and more difficulty in close relationships ($r(75) = 0.59, p < .01$). Surprisingly, participant initiation of negative social interactions was related to the experience of positive social interactions initiated by others, $r(75) = 0.40, p < .01$. This may be further evidence that greater overall engagement in social activities, whether that engagement is negative or positive, creates more opportunities for interaction initiations from others, both positive and negative. This may also be indicative of attempts by others to placate borderline individuals after they have done something negative, or to otherwise make positive repairs in the wake of negative interactions.

**Positive interactions initiated by others.** Participants reported on how frequently other individuals initiated positive social interactions with them. For example, participants reported how frequently they received compliments from others. The frequency with which they reported having experienced positive social engagement initiated by others was unrelated to most measures of social functioning. Two exceptions emerged, as discussed earlier: 1) participants
who reported that others acted positively toward them more often also reported that they themselves engaged in more antisocial acts; and 2) participants who reported that others acted positively toward them more often also reported that they themselves initiated both more positive and more negative social interactions toward others. These findings might indicate that some individuals are more socially engaged than others and generally have more opportunity for both negative and positive social contact. Frequency of positive interactions initiated by others was unrelated to negative interactions initiated by others ($r(75) = 0.16, p > .05$), overall self-reported satisfaction with social life ($r(75) = -0.18, p > .05$), overall social functioning on the SFQ ($r(75) = -0.70, p > .05$) and the SAS ($r(67) = 0.22, p > .05$), and experiences in close relationships ($r(75) = -0.21, p > .05$). Initiation of positive interactions may be an index, at least in part, of greater overall social engagement. Frequent social engagement, however, does not necessarily result in positive outcomes or overall functioning in life roles.

Measures of general social functioning were largely intercorrelated. Lower general social satisfaction was associated with lower functioning on the SFQ ($r(75) = 0.62, p < .01$) and the SAS ($r(67) = 0.34, p < .01$), as well as poorer experiences in close relationships ($r(75) = 0.62, p > .05$). SFQ approximation of overall social functioning were associated as expected with approximation of social functioning on the SAS ($r(67) = 0.59, p < .01$). Finally, more difficulty in close relationships was associated with poorer social functioning on both the SFQ ($r(75) = 0.76, p < .01$) and the SAS ($r(67) = 0.51, p < .01$).
Hypothesis 1

The borderline group will report lower social functioning across all measures. Similarly, borderline features will be associated with lower social functioning across all measures.

In comparison to the control group, the borderline features group reported lower social functioning on most measures. This group reported engaging in more antisocial acts ($t(75) = 2.77, p = 0.007, R^2 = 0.093$), initiating more negative interactions with others ($t(75) = 5.41, p = 0.000, R^2 = .281$), being the recipient of more negative interactions initiations from others ($t(75) = 4.81, p = 0.000, R^2 = 0.236$), lower global social satisfaction ($t(75) = 3.13, p = 0.002, R^2 = 0.116$), worse overall functioning on the SFQ ($t(75) = 5.72, p = 0.000, R^2 = 0.304$), and poorer experiences in close relationships ($t(75) = 4.74, p = 0.000, R^2 = 0.230$). On the SAS, they reported worse overall functioning ($t(67) = 3.17, p = 0.002, R^2 = 0.130$) as well as poorer domain-specific functioning in areas of work ($t(35) = 2.11, p = 0.042, R^2 = 0.113$), school ($t(31) = 2.71, p = 0.010, R^2 = 0.194$), leisure ($t(67) = 1.87, p = 0.066, R^2 = 0.050$), and interactions with extended family ($t(67) = 2.71, p = 0.004, R^2 = 0.118$).

Contrary to expectations, the borderline features and control groups did not differ on commission of prosocial acts ($t(75) = 0.01, p = 0.993, R^2 = 0.000$). Individuals in the borderline features group reported that they engaged in prosocial acts at similar rates to the control group.
Failure to find a significant difference between groups in this regard is interesting in light of the chronic negative affect that characterizes BPD, as well as significant stigma that paints these individuals as difficult, manipulative, and “bad.” Borderline and control groups also did not differ on initiation of positive social interactions ($t(75) = 0.24, p = 0.810, R^2 = 0.001$), experience of initiation of positive interactions from others ($t(75) = 0.46, p = 0.646, R^2 = 0.003$), functioning in the domain of housework ($t(31) = 1.58, p = 0.125, R^2 = 0.074$), partner role functioning ($t(14) = 0.09, p = 0.929, R^2 = 0.001$), or functioning within the close family unit (including those with whom the participant lives; $t(67) = 1.64, p = 0.106, R^2 = 0.039$). Importantly, only a subset of individuals endorsed that housework ($n = 33$) and partner roles ($n = 16$) were important domains of functioning for them. This limits the sample size for these group comparisons. When a role is not endorsed on the SAS, it is not rated. Table 6.3 presents additional information, including group means.
Association of borderline features with social functioning was further assessed dimensionally. Full dimensional results are reported in Table 6.4. Notably, results are consistent across measures of borderline pathology. T scores on the SNAP-2 and the PAI-BOR, as well as number of diagnostic criteria endorsed on the SCID-II show the same pattern of association with all social functioning measures. Individuals who endorsed greater borderline pathology on these
measures reported committing more mild antisocial acts, initiating more negative social interactions, experiencing more negative social interactions initiated by others, lower leisure satisfaction, poorer experiences in close relationships, and lower overall social functioning on both the SFQ and the SAS.

Hypothesis 2

A tendency to rate others as less trustworthy will be related to poorer social functioning for the sample overall.

Average raw trustworthiness appraisals were not associated with most measures of social functioning. However, they were associated with prosocial acts on the BNQ, such that individuals who rated faces are more trustworthy reported engaging in more prosocial acts, consistent with positive approach behavior, $r(75) = 0.26, p < 0.05$. Average raw trustworthiness appraisals were also associated with better overall social functioning on the SFQ, $r(75) = -0.25, p < 0.05$. Using a two-tailed test, the association of average raw trustworthiness appraisals was also

Table 6.4. Correlations of social functioning scales with measures of borderline features.

<table>
<thead>
<tr>
<th></th>
<th>SNAP-2</th>
<th>SCID-II</th>
<th>PAI-BOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNQ (Prosocial)</td>
<td>-.060</td>
<td>.147</td>
<td>.055</td>
</tr>
<tr>
<td>NPQ (Antisocial)</td>
<td>.425**</td>
<td>.504**</td>
<td>.500**</td>
</tr>
<tr>
<td>SIIS: Participant Positive</td>
<td>.018</td>
<td>-.136</td>
<td>.145</td>
</tr>
<tr>
<td>SIIS: Participant Negative</td>
<td>.682**</td>
<td>.493**</td>
<td>.583**</td>
</tr>
<tr>
<td>SIIS: Other Positive</td>
<td>.101</td>
<td>.080</td>
<td>.031</td>
</tr>
<tr>
<td>SIIS: Other Negative</td>
<td>.583**</td>
<td>.500**</td>
<td>.621**</td>
</tr>
<tr>
<td>Leisure Satisfaction</td>
<td>.438**</td>
<td>.411**</td>
<td>.456**</td>
</tr>
<tr>
<td>SFQ</td>
<td>.650**</td>
<td>.586**</td>
<td>.575**</td>
</tr>
<tr>
<td>SAS: Total</td>
<td>.496**</td>
<td>.397**</td>
<td>.622*</td>
</tr>
<tr>
<td>ECR</td>
<td>.615**</td>
<td>.582**</td>
<td>.624**</td>
</tr>
</tbody>
</table>

Note. T scores are used for SNAP-2 and PAI-BOR

** $p < .01$, * $p < .05$ (two-tailed tests)
trending toward significant for overall social functioning as measured by the SAS, $r(67) = -0.22$, $p < 0.1$. This association was significant when using a one-tailed test, which was justified in this case by the *a priori* hypothesis that lower trustworthiness ratings would predict poorer social functioning. Individuals who rated faces as more trustworthy on average also reported better social functioning. Correlations of social functioning scales with average trustworthiness appraisals and priming influences are presented in Table 6.5.

Table 6.5. Correlations of social functioning scales with average trustworthiness appraisals and priming influence in the overall sample.

<table>
<thead>
<tr>
<th></th>
<th>Average Trustworthiness Appraisal</th>
<th>Negative Difference</th>
<th>Positive Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNQ (Prosocial)</td>
<td>.256*</td>
<td>-.073</td>
<td>.208</td>
</tr>
<tr>
<td>NPQ (Antisocial)</td>
<td>-.085</td>
<td>.046</td>
<td>-.136</td>
</tr>
<tr>
<td>SIIS: Participant Positive</td>
<td>.110</td>
<td>-.081</td>
<td>.046</td>
</tr>
<tr>
<td>SIIS: Participant Negative</td>
<td>-.044</td>
<td>.011</td>
<td>-.066</td>
</tr>
<tr>
<td>SIIS: Other Positive</td>
<td>.110</td>
<td>-.051</td>
<td>.050</td>
</tr>
<tr>
<td>SIIS: Other Negative</td>
<td>-.136</td>
<td>.035</td>
<td>-.207</td>
</tr>
<tr>
<td>Leisure Satisfaction</td>
<td>.113</td>
<td>.008</td>
<td>-.177</td>
</tr>
<tr>
<td>SFQ</td>
<td>-.250*</td>
<td>.013</td>
<td>-.160</td>
</tr>
<tr>
<td>SAS: Total</td>
<td>-.218^</td>
<td>-.118</td>
<td>.152</td>
</tr>
<tr>
<td>ECR</td>
<td>-.218</td>
<td>.040</td>
<td>-.119</td>
</tr>
</tbody>
</table>

** $p < .01$, * $p < .05$, ^ $p < .1$ (two-tailed tests)

**Hypothesis 3**

*Trustworthiness ratings will mediate the relationship between borderline features and social functioning.*

To test hypothesis 3, I first examined the association of trustworthiness appraisals and priming effects with social functioning specifically within the borderline features group. These results are generally consistent with the associations observed within the overall sample,
described above. Two differences emerged: average trustworthiness appraisals were not associated with overall social functioning on the SFQ and SAS for the borderline group as was the case for the overall sample. Additionally, in the borderline group but not the overall sample, positive difference scores were associated with the experience of negative social interactions initiated by others. Individuals in this group who were more susceptible to the influence of positive primes reported having fewer negative social interactions initiated by others, \( r(75) = -0.38, p < 0.05 \). One speculative explanation for this finding is that for individuals in the borderline group, greater influence of positive information is protective against perceiving interactions as negative. Table 6.6 presents further information, including correlations of trustworthiness appraisals and priming effects with measures of social functioning.

<table>
<thead>
<tr>
<th></th>
<th>Average Trustworthiness Appraisal</th>
<th>Negative Difference</th>
<th>Positive Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNQ (Prosocial)</td>
<td>.387*</td>
<td>-.137</td>
<td>.101</td>
</tr>
<tr>
<td>NPQ (Antisocial)</td>
<td>.026</td>
<td>.197</td>
<td>-.261</td>
</tr>
<tr>
<td>SIIS: Participant Positive</td>
<td>.209</td>
<td>-.069</td>
<td>.129</td>
</tr>
<tr>
<td>SIIS: Participant Negative</td>
<td>.164</td>
<td>.329</td>
<td>-.217</td>
</tr>
<tr>
<td>SIIS: Other Positive</td>
<td>.284</td>
<td>-.066</td>
<td>.114</td>
</tr>
<tr>
<td>SIIS: Other Negative</td>
<td>.050</td>
<td>.275</td>
<td>-.380*</td>
</tr>
<tr>
<td>Leisure Satisfaction</td>
<td>-.177</td>
<td>.110</td>
<td>-.298</td>
</tr>
<tr>
<td>SFQ</td>
<td>-.082</td>
<td>.260</td>
<td>-.338</td>
</tr>
<tr>
<td>SAS: Total</td>
<td>-.031</td>
<td>-.174</td>
<td>.252</td>
</tr>
<tr>
<td>ECR</td>
<td>-.262</td>
<td>.267</td>
<td>-.299</td>
</tr>
</tbody>
</table>

**p < .01, *p < .05 (two-tailed tests)**

The present study included multiple measures of borderline features and social functioning. To test whether trustworthiness appraisals mediated the relationship between
borderline features and social functioning, I used the SNAP-2 for borderline features to be consistent with the manner in which groups were defined. Regression analysis were performed only for those areas of social functioning that were associated with average trustworthiness appraisals for the overall sample (prosocial behaviors, SFQ, SAS; included in Table 6.5). This yielded three individual mediation analyses.

**Prosocial behaviors.** Average trustworthiness appraisal scores were regressed on borderline features (number of criteria endorsed on SNAP-2). This revealed a significant contribution of borderline features ($\beta = -0.283, p = .013$). A second regression included borderline features as a predictor variable and prosocial behaviors as the outcome variable. Consistent with Table 6.4, this was not significant ($\beta = -0.41, p = 0.724$). Borderline features did not predict prosocial acts. Mediation analysis stopped here.

**SFQ scores.** Second, I ran a regression analysis with borderline features as a predictor variable and SFQ score as the outcome variable. Consistent with Table 6.4, this was significant ($\beta = -0.64, p < 0.001$). Borderline features significantly predicted poor social functioning as measured on the SFQ. Again, I regressed average trustworthiness appraisal scores on borderline features (number of criteria endorsed on SNAP-2). This revealed a significant contribution of borderline features ($\beta = -0.283, p = .013$) to trustworthiness appraisal. I next regressed both borderline features and average trustworthiness appraisals on SFQ scores in the same model. In this model, average trustworthiness appraisal did not predict SFQ scores ($\beta = -0.074, p = 0.427$). Average trustworthiness appraisals were not a plausible mediator of the association between borderline features and SFQ scores.

**SAS scores.** I repeated the same process using SAS scores as the dependent variable of interest. Again, I regressed average trustworthiness appraisal scores on borderline features.
(number of criteria endorsed on SNAP-2). This revealed a significant contribution of borderline features ($\beta = -0.283, p = .013$) to trustworthiness appraisal. I then ran a regression analysis with borderline features as a predictor variable and SAS score as the outcome variable. Consistent with Table 6.4, this was significant ($\beta = 0.481, p < 0.001$). Borderline features significantly predicted poor social functioning as measured on the SAS. I next regressed both borderline features and average trustworthiness appraisals on SAS scores in the same model. In this model, average trustworthiness appraisal did not predict SAS scores ($\beta = -0.045, p = 0.701$). This indicates that average trustworthiness appraisals are not a plausible mediator of the association between borderline features and SAS scores. Hypothesis 3 is disconfirmed.

**Hypothesis 4**

*The magnitude of negative affective priming influence on trustworthiness appraisals will predict poorer social functioning for the sample overall.*

Evidence does not support hypothesis 4. Negative (and positive) difference scores were not related to any measures of social functioning. As described in Chapter 3, these difference scores are an index of the effect of priming. Higher difference scores indicate greater influence of either negative or positive primes on trustworthiness ratings, relative to overall trustworthiness ratings for each individual. These scores essentially control for within-subject variability in average or baseline trustworthiness appraisals to capture the specific influence of affective primes. Table 6.4 includes correlations between negative difference scores and social functioning measures.
Hypothesis 5

*Negative affective priming influence will mediate the relationship between borderline features and social functioning.*

Negative differences scores, an index of the influence of negative affective primes, were not related to social functioning measures. As outlined in Tables 6.5 and 6.6, this is true when considering the sample overall and when considering the borderline features group specifically. Thus negative affective priming influence is not viable as a mediator of the association between borderline features and social functioning. Mediation analysis is not warranted, and hypothesis 5 is disconfirmed.

Discussion

Here I examined whether trustworthiness appraisals are related to social functioning for the sample overall and for individuals in the borderline features group specifically. I also examined whether the influence of affective primes predicted social functioning across an array of measures. Four main findings resulted.

First, newly developed measures described in Chapter 5, including the SIIS, the NPQ, and the BNQ, generally perform as expected. However, it appears that overall the negatively oriented parts of these measures are most consistently related to poor social functioning. For example, the commission of antisocial but not prosocial acts is broadly related to poor social functioning. Additionally, participant- and other-initiated negative interaction frequency are broadly associated with poorer social functioning, but participant- and other-initiated positive interactions are not consistently related to other measures of social functioning. These results are consistent with the idea that negative behavior may be most central to whether an individual is functioning well in his or her social environment.
Second, the borderline features group endorsed poorer social functioning than the control group across an array of measures. A few exceptions emerged on measures of positive social engagement: the borderline group did not differ from the control group on prosocial acts (on the BNQ) or participant- or other-initiated positive interactions (on the SIIS). Individuals with borderline features appear to engage in positive interactions and behaviors at a level comparable with the control group. However, they engage more often in negative interactions and behaviors and endorse lower overall social functioning. This is further evidence of the centrality of negative processes in borderline personality. Together, these findings suggest that negative processes are central for borderline personality and that responses made by the borderline group were likely not broadly affected by a general negative response bias—some measures show evidence of positive behaviors and interactions comparable to the control group.

Third, average trustworthiness appraisals were associated with measures of overall social functioning (the SFQ and SAS) in the total sample. However, this association was not significant when considered specifically in the context of the borderline features group. Consistent with mediation analyses conducted here, this suggests that biased trustworthiness appraisals are not a plausible mechanism by which borderline features give rise to poorer social functioning.

Fourth, degree of priming influence was not associated with social functioning. In other words, the degree to which negative or positive affective information influenced trustworthiness appraisals was not associated with how well individuals were functioning within their social environments. We had expected that greater influence of negative information on the social cognitive process of ascribing trustworthiness to another individual would predict poorer social functioning. This was not the case.
Overall, results indicate that despite its reliability and replicability, the finding that individuals with borderline features have biased trustworthiness appraisal may not have much practical significance. Although individuals with borderline features rate others as less trustworthy and act accordingly in laboratory situations, this bias may not translate to real-world social functioning. Further work is needed to examine whether trustworthiness bias is observed in more ecologically valid tasks. Bias that emerges in more ecologically valid paradigms may have more external predictive validity. Furthermore, although negative affective information had a greater influence on trustworthiness appraisal for individuals with borderline features, this influence was not predictive of social functioning. This suggests that although individuals with borderline features may make social judgments that are unduly influenced by negative emotional information, this influence is not a primary process by which these individuals come to have poor social functioning. The measurement of affective priming influences in more ecologically valid paradigms may yield different results and is a warranted target for future research.
Chapter 7: General Discussion and Conclusions
A growing body of literature shows that individuals with BPD and subclinical borderline features rate others as less trustworthy and act accordingly. The goal of this dissertation was to replicate and extend these findings. Here I sought to extend this work by exploring the impact of affective information on the social cognitive process of appraising trust. I also sought to determine whether perceptions of low trustworthiness influence recognition of previously viewed faces. Finally, I sought to determine whether trust processing biases observed in the laboratory are associated with real-world social functioning. Participants were divided into two groups on the basis of borderline traits endorsed on the SNAP-2. Individuals who endorsed 2 or fewer traits were placed in a control group, while individuals who endorsed 3 or more traits were placed into a borderline or borderline features group. I hypothesized that the borderline features group would be more susceptible to affective priming when making trustworthiness judgments. Additionally, I hypothesized that the borderline group would make more errors and would be more influenced by perceptions of low trustworthiness during a recognition task. Finally, I predicted that low trustworthiness ratings and susceptibility to affective priming would predict poorer social functioning, particularly for the borderline group.

**Brief Summary of Primary Results**

Individuals in the borderline features group rated faces as less trustworthy than the control group overall. The borderline features group appraised faces as less trustworthy following negative, neutral, and positive primes. This group also showed evidence of greater susceptibility to affective priming. Trustworthiness ratings made by the borderline group were significantly lower following negative affective primes relative to the control group. Positive affective primes did not differentially influence ratings made by the two groups. When examined dimensionally, these results were broadly consistent. For the entire sample, greater number of
borderline symptoms endorsed predicted lower trustworthiness ratings and influence of negative affective primes. The affective instability subscale of the PAI showed significant association with trustworthiness ratings and influence of negative affective primes. Other subscales of the PAI borderline scale generally did not show association with either average trustworthiness appraisals or priming influences. Contrary to expectations and previous findings, rejection sensitivity and childhood trauma were not related to trustworthiness ratings or to the influence of affective primes in this sample.

On the recognition task that followed the trustworthiness appraisal task after a short delay, the borderline group again rated faces as less trustworthy than the control group. The control group expressed greater certainty about their recognition judgments. This was true for both correct and incorrect responses. Groups did not differ on accuracy, which was quite high overall. Participants made more “miss” errors than “false positive” errors. The borderline features group showed a trend to rate faces to which they responded incorrectly as less trustworthy than faces to which they responded correctly. This difference was not found in the control group.

Contrary to hypotheses, trustworthiness bias and susceptibility to negative affective priming were largely unrelated to social functioning across several domains and measurement tools. This was true of the sample overall. This was also true when groups were considered separately: trustworthiness appraisal and priming effects were not related to social functioning for individuals within only the borderline features group or within only the control group.

Consistent with the notion that BPD is associated with—and indeed characterized by—interpersonal dysfunction, the borderline group reported poorer social functioning on the majority of measures included.
Interpretations and Implications of Findings

Bias and the Influence of Affective Priming

These findings are consistent with past work showing a response bias for individuals with BPD and subclinical borderline features to rate others as less trustworthy. Taken together with previous findings the current work suggests that trustworthiness appraisal bias is a reliable and robust social cognitive disturbance for individuals with BPD or borderline features. It is unclear from the current findings, however, whether this disturbance is best conceptualized as a response bias or an interpretive bias. BPD is consistently associated with negative affect and negative bias on a broad range of tasks. It is possible in principle that individuals in the borderline features group simply rated faces as more negative regardless of their interpretation of those faces. If this was the case, we might also expect them to rate clearly unambiguous neutral information (e.g., simple shapes) as more negative. However, the faces used in the task were inherently ambiguous in terms of trustworthiness and emotional expression (in some cases). Thus it may be that active interpretation of the faces yielded lower ratings of trustworthiness for the borderline features group. If this were the case, then we would expect this group to rate clearly neutral information more negatively. Disambiguation of response bias versus interpretation bias may be a useful target of future work.

Regardless of whether we conceptualize the disturbance observed here as a response bias or an interpretive bias, it is clear that it is influenced by affective information. The bias seen here was present across priming conditions—the borderline group consistently rated faces as less trustworthy across neutral, positive, and negative priming conditions. However, the bias was significantly greater for individuals with borderline features compared to controls after the presentation of specifically negative affective information. To the author’s best knowledge, this
is the first study to show that compared to controls, trustworthiness appraisals made by individuals with borderline features are more greatly influenced by negative affective information. Individuals with borderline features are thus similar to individuals with schizophrenia in this regard (see Hooker et al., 2011). This is consistent with the frequent manifestation of paranoia in both disorders: in schizophrenia paranoia is typically pervasive and stable. In borderline personality disorder, paranoia is stress-related and transient. The influence of negative affective primes was not attributable to trait anxiety, social anxiety, or state changes in anxiety or mood. Negative affective information had greater influence over trustworthiness appraisals for individuals in the borderline features group relative to the control group. The influence of positive affective information on trustworthiness appraisals did not differ between groups. Individuals with borderline features were differentially and specifically influenced by negative affective information.

One potential mechanism underlying the association of borderline features and biased trustworthiness appraisal is rejection sensitivity. Past work has shown that rejection sensitivity mediates the association of borderline features and trust appraisals (Miano et al., 2013). However, the results of the current study are inconsistent with this finding: rejection sensitivity was not related to trustworthiness appraisals or to priming effects. This was true when assessed using the same measure as Miano et al. (2013), and when assessed using an additional measure. This puzzling difference may be the result of differences in sample characteristics. For example, it is possible that the sample used by Miano et al. (2013) was more severe, but this is unlikely given that they too examined individuals with borderline features who did not necessarily meet full diagnostic criteria for BPD. The reverse is more likely true, as Miano et al. sampled undergraduates without an explicit attempt to capture individuals with borderline pathology.
Because Miano et al.’s sample was younger (mean age = 19.8) than the current sample, it is also possible that young adults in college environments may be differentially affected by rejection sensitivity than older adults or young adults not in college. The current study raises questions about the plausibility of rejection sensitivity as a mechanism for biased trust appraisal in BPD.

As predicted, the impact of negative affective primes was uniquely predicted by the affective instability subscale of the broader PAI Borderline scale. This subscale was associated with negative priming effects (negative difference scores), as well as appraisals made after negative primes and neutral primes. Association between PAI affective instability and influence of negative priming effects provides some validation that the priming effects witnessed here are likely tapping into underlying affective instability for participants with borderline features (and participants more broadly). Furthermore, these findings suggest that affective instability may be a mechanism for biased or negative interpretations of social information. Biased or negative interpretations of social information may in turn drive interpersonal dysfunction. Although the current study did not find a connection between affective priming influence and social functioning, it remains plausible that affective instability may be one mechanism that supports disturbed social cognition in BPD.

The affective instability subscale of the PAI was not associated with ratings made after positive primes, providing further evidence that positive and negative emotional processing in BPD may rely on separable mechanisms that influence and are influenced by differentiated systems. The PAI identity disturbance subscale was also related to appraisals made after negative primes, but not to negative difference scores/negative priming effects. Overall, it seems that the exaggerated influence of negative affective information that is indexed on the behavioral task used here validly captures affective instability as reported on the PAI. Individuals who
themselves report greater difficulty regulating affect show behavioral evidence indicative of the
same. Interestingly, no trustworthiness ratings (regardless of priming condition) were related to
the negative relationships subscale of the PAI. This is consistent findings from the current study
showing that trustworthiness appraisals and priming influence were not related to measures of
social functioning.

Like rejection sensitivity, childhood trauma also did not predict biased trustworthiness
appraisals, nor did it not moderate the relation of borderline features and biased trustworthiness
ratings. This is somewhat surprising given past work that suggests a direct link between early
aversive experiences and trust appraisal. One possible explanation for the findings is that
childhood trauma may not have been accurately or consistently reported, at least for the control
group. The CTQ includes a scale that captures minimization/denial of early childhood adversity.
The control group had high rates of minimization/denial according to this scale. This may render
the results uninterpretable. Alternatively, it may be the case that early childhood trauma makes
some individuals more trusting and some individuals less trusting. In other words, it may
influence trust processing differentially depending on the individual and other contextual
variables. Other work has suggested that betrayal traumas produce a positive trustworthiness bias
in some individuals such that they become too trusting of others (Gobin & Freyd, 2009;
Zurbriggen & Freyd, 2004). If some individuals with childhood trauma become more trusting
and some individuals become less trusting, then this may obscure linear estimates of the
association between trauma and trust. Further work is needed to better understand what causes
these two divergent consequences following childhood trauma. Additionally, the current sample
may also reflect a lower base rate of childhood abuse than other samples, which may truncate
variance and make the results difficult to interpret. The borderline group did not differ from the
control group on several subscales of the Childhood Trauma Questionnaire, including physical abuse, sexual abuse, and physical neglect. This is potentially consistent with the manner in which groups were defined. Given that the borderline group was defined based on features rather than full diagnosis, it may be a less severe group than is typically observed in clinical settings. Some of the work (e.g., Zanarini, 2000) that has shown a link between BPD and early trauma has focused on severe, inpatient samples. Childhood trauma may not be a part of the etiological picture for many individuals with BPD. Emotional vulnerability coupled with a poor but non-abusive fit with parenting styles and early environments may be enough to give rise to BPD (Linehan, 1993). Finally, it may be the case that childhood trauma influences other aspects of social cognition and trustworthiness appraisals.

**Face Recognition and Trustworthiness Appraisal**

Average trustworthiness ratings made after the affective priming task were highly correlated for all participants with average trustworthiness appraisals made during the recognition. This suggests that trustworthiness appraisals are relatively stable over a short delay period, and further confirms that *overall, average* appraisals are unlikely to be driven primarily by momentary mood or anxiety. Importantly, however, this does not mean that trustworthiness appraisals are entirely stable for individuals with BPD or borderline features. Here we saw that these appraisals were influenced by negative affective information. This means that these social judgments may be malleable based on affective context, whether that context is internal (e.g., momentary changes in affect) or external (e.g., negative affective information in the environment). *Overall, on average,* ratings were consistent across time. However, within a given time point (during the affective priming task), individuals with borderline features were significantly influenced by negative affective information on a moment-to-moment basis.
Individuals with borderline features did not differ from the control group on overall face recognition accuracy. The lack of differences between groups may be attributable to the relatively short delay time. It may also be attributable to the manner in which groups were defined. The borderline group is best regarded as representing individuals with borderline features who do not necessarily meet full diagnostic criteria for BPD. Memory differences between this group of individuals with borderline features rather than full diagnosis and individuals in the control group may be attenuated relative to past studies that have compared individuals on the basis of diagnostic status. The borderline group in this study also had a higher estimated IQ than the control group, and might reasonably be expected to perform better on tests of neurocognition.

Although the borderline group did not show a difference in accuracy of face recognition, an important finding emerged with regard to trustworthiness appraisal and accuracy. The borderline group rated faces to which they responded incorrectly as less trustworthy compared to faces to which they responded correctly. In other words, incorrect responses were associated with lower trustworthiness ratings than correct responses for the borderline group. The control group did not show any similar association between trustworthiness and response accuracy. This finding provides further evidence that individuals with borderline features are more susceptible to negative emotional information. Cues of low trustworthiness (whether they are objectively present or perceived owing to interpretation bias) are fundamentally a type of negative social information. Untrustworthy individuals are more likely to initiate betrayal, rejection, or other aversive social experiences. Evidence now converges to show that negative emotional or social information differentially influences cognitive processes for individuals with BPD. The findings
of the current study suggest that negative information also has a differential impact on cognitive processes in the context of a socially relevant task.

The finding that individuals in the borderline group made lower trustworthiness appraisals for faces to which they responded incorrectly may also be considered an example of a source monitoring error. Perceiving a face as less trustworthy may lead individuals with borderline features to conclude that they must have seen it before. Although source memory seems to be intact overall in BPD, lower source memory is associated with greater hostility and suspiciousness for individuals with the disorder, even controlling for negative affect (Minzenberg et al., 2006). This means that individuals with BPD who are high on hostility and suspiciousness have greater difficulty monitoring the initial source of their information—whether that information was actually perceived or whether it was imagined. Further work may do well to examine how personal hostility and suspiciousness interact with social judgments to predict source monitoring errors. It may be the case that the perception of others as less trustworthy confuses some individuals, but not others, with borderline features to believe that they have seen the face before.

Finally, individuals with borderline features expressed less certainty about their recognition judgments overall than individuals in the control group. In contrast to our expectations, these individuals did not express greater certainty than the control group about incorrect responses. This is seemingly inconsistent with ideas about hyper-mentalizing in BPD. Perhaps promisingly for individuals with borderline pathology, this finding suggests some awareness of the difficulty of remembering this social information. Although individuals with borderline features may be more influenced by negative social information as they attempt to recall faces, they are also seemingly aware of their own fallibility.
Trustworthiness Appraisal and Social Functioning

The most surprising result of the current study is the lack of association between the social cognitive process of trustworthiness appraisal and self-reported social functioning. Broadly, trustworthiness appraisals and affective priming influence were unrelated to real-world social/interpersonal functioning across a variety of domains and scales.

Biased trustworthiness appraisal manifests behaviorally in laboratory settings. Not only do individuals with borderline pathology rate others as less trustworthy—they also act accordingly. This results in poorer outcomes on tasks that require social exchange for optimal performance. The current findings suggest that this may not be the case in real-world situations. Further work is needed to examine whether trustworthiness appraisal is biased for individuals with borderline pathology in more ecologically valid paradigms, and whether bias that emerges in these paradigms has greater external predictive validity. In particular, further work is needed to determine whether and how trustworthiness biases might differ in different kinds of relationships. Thus far, trustworthiness biases have been demonstrated exclusively in tasks that require an individual to make judgments about a novel other. Upon the very first viewing of a face or during the first interaction with another person, individuals with BPD judge other individuals as less trustworthy. It is unclear whether trustworthiness biases would be consistent in situations in which participants are asked to appraise the behavior or trustworthiness of individuals with whom they are more familiar. The finding that laboratory-produced trustworthiness bias was not predictive of social functioning in the current study might be reflective of the face that much of social functioning relies on relationships with known individuals rather than single interactions with new people. Trustworthiness biases that may exist in the context of more established relationships might be more predictive of social functioning.
Additionally, we had expected that the degree to which trustworthiness appraisals were influenced by negative affective information would be predictive of poor social functioning. Individuals who revise social judgments more readily based on negative emotional information may engage in fewer social activities, perceive social interactions more negatively, and have more difficulty functioning in social roles. Findings from the current study are not consistent with this expectation. This suggests that although individuals with borderline features may make social judgments that are unduly influenced by negative emotional information, this influence is not a primary process by which these individuals come to have poor social functioning. Although the influence of emotional information on the social cognitive process of trait appraisal may be an important process to understand in terms of elucidating borderline pathology, it may not have great practical significance. Again, the measurement of affective priming influences in more ecologically valid paradigms may yield different results and is a warranted target for future research.

**Limitations**

The current study may be limited owing to several aspects of the sample. First, the sample size is relatively small. However, many predicted effects reached statistical significance even with the low power that accompanies small sample sizes. Additionally, the majority of null findings, particularly those for the connection between trustworthiness appraisal processes and social functioning, are not in the realm of trending toward significance and would be unlikely to be significant even in a much larger sample. Additionally, the sample is predominantly female. This is unlikely to represent a true difference in the base rate for the disorder. Although the DSM continues to suggest that BPD occurs at a rate three times higher in women than in men, well-controlled studies suggest that BPD base rates for men and women are roughly equivalent (e.g.,
Torgersen, Kringlen, & Cramer, 2001). The sample is also of high average intelligence and the borderline group scored significantly higher on estimated IQ. This is consistent with the generally high education level of this sample, but not consistent with the typical patient presentation and may limit generalizability. This difference in estimated IQ might also help explain why the two groups did not differ on accuracy of face recognition in Chapter 4.

The group differentiation in the current study is not based upon the diagnostic cutoff for BPD (5/9 symptoms), but rather a subclinical cutoff (3/9 symptoms). This is consistent with the Collaborative Longitudinal Personality Disorder Study (CLPS; see Gunderson et al., 2011), one of the largest ongoing studies of BPD course. The CLPS considers patients in remission when they meet for 2 or fewer BPD criteria. The control group used in the current study was defined by the presence of 2 or fewer BPD symptoms as measured using the SNAP-2. Individuals in the CLPS who meet for 3 or more BPD criteria are considered not to have attained remission and to have significant levels of borderline pathology. Consistent with this cutoff, we found significant differences in how individuals who endorse 3 or more BPD criteria appraise trustworthiness. We also found other significant group differences indicating that setting a threshold at 3 or more BPD symptoms captures significant pathology. Still, the cutoff used here is not a diagnostic cutoff, and generalization to clinical populations should be made cautiously.

Borderline personality disorder is highly comorbid with mood disorders, so much so some argue that BPD may be best regarded as a mood disorder itself (Akiskal, 2002; Sauer-Zavala & Barlow, 2014). In one study, concurrent major depression was found in 61% of individuals with BPD, while dysthymia and bipolar disorder were also highly comorbid (12% and 20%; Zimmerman & Mattia, 1999). Over the course of a lifetime, the majority of individuals with BPD will have a major depressive episode (83%; Tomko et al., 2014). However, BPD is
also highly comorbid with other disorders, including anxiety disorders, substance use disorders, PTSD, eating disorders, and other personality disorders (Grant et al., 2008; Zimmerman & Mattia, 1999; Lenzenweger et al., 2007; McGlashan et al., 2000). Patients with BPD are generally more likely to have comorbid disorders than patients who have other psychiatric disorders but not BPD (see Gunderson, 2001). The notion that BPD is best considered a mood disorder is not widely supported (Choi-Kain & Gunderson, 2015). However, high rates of mood disorder in BPD generally and in this sample specifically (e.g., the borderline features group endorsed higher rates of mood disorders and high scores on the BDI-II) must be considered when interpreting the current results. As is often the case with studies of BPD, it is difficult to be certain that the findings here are due to BPD and not attributable at least in part to coexisting depression. In order to better differentiate what effects are due to BPD and which are due to depression or other mood disorder, a psychiatric control condition might be helpful (e.g., a group with dysthymia/persistent depressive disorder, as used by Hooley et al., 2010). The current study has not attempted to “parcel out” the effects of depression. Statistical controls for depression would not be appropriate given that the groups here were not randomly assigned, and that higher levels of depressive symptoms may be a normal part of the clinical manifestation of BPD (see Korfine & Hooley, 2000). Nevertheless, future work with psychiatric control groups would be well poised to examine the relative impact of BPD and depression on social cognitive processes like those examined here.

Finally, this study was not conducted in a manner that left researchers fully blind to group. Diagnostic interviews were conducted prior to viewing SNAP-2 results, which limits the likelihood that they were influenced by factors other than participant report. However, participants responded to a variety of advertisements placed in the community, and frequently
told researchers (the current author and research assistants) that they were interested in participating either because they had BPD or because they felt they were mentally healthy. It is possible that this knowledge influenced test administration. However, the behavioral tasks were unlikely to be biased by researchers as they were administered via computer. Other assessment procedures were conducted according to available guidelines for standard administration.

**Future Directions**

The current study focused on trustworthiness appraisals made during a controlled laboratory task. This work extends previous work that has also studied trustworthiness appraisals in laboratory conditions. Further work is needed to examine how individuals with and without borderline pathology make trustworthiness judgments in tasks or situations with greater ecological validity. Trustworthiness appraisal biases in individuals with BPD and borderline features are produced reliably in laboratory settings, but is currently unclear whether these biases manifest in real-time social interactions, and if bias manifests similarly in laboratory versus natural or more ecologically valid settings.

Natural interactions between people often produce a variety of facial affect. The current study used faces with mostly neutral expressions, and was not designed to allow for comparisons to be made between faces of various emotional expression or intensity. Currently we know of ongoing but unpublished work by Miano and colleagues that investigates the role of emotional expression and emotion recognition in trustworthiness processing for individuals with BPD. This work is essential to integrating research on emotion recognition and trait appraisal. Although these processes both fall squarely within the realm of social cognition, their overlap or association has been largely ignored thus far.
Also like other studies of trustworthiness appraisal in BPD, the current study required individuals to make fairly rapid decisions about trustworthiness. Future research might expand upon findings from studies using these types of paradigms to begin to examine how individuals with BPD make trustworthiness judgments when given more time and more information. The result of deliberate decision making may differ from what is produced in paradigms that may elicit automatic or implicit reactions to faces. In the process of slowing this process down to include deliberate decision making, future research would also be well positioned to examine the flexibility of trustworthiness appraisals made by individuals with BPD. For example, are individuals with BPD less able to integrate new information to change previously made judgments? We are currently pursuing this question in other work.

Future work is needed to understand why negative affective information seems to exert undue influence on trustworthiness appraisal in borderline pathology, as well as whether this influence extends to other aspects of social cognition. One possible mechanism for undue affective influence is problems with cognitive control of this information. Dysfunctional regulatory control over an over-reactive amygdala may explain the affective priming effect seen here. Work that integrated neuroimaging with this type of behavioral task may prove fruitful for better understanding social cognitive biases in BPD.

Finally, these findings may be beneficially incorporated into future treatment research. A recent promising study by Goodman et al. (2014) showed that dialectical behavior therapy (DBT) may improve emotion regulation via improved amygdala habituation. This suggests a possible mechanism by which this well-regarded treatment may prove effective for individuals with BPD. Additionally, it suggests that we might expect changes in the influence of affective information on social cognitive processes like those examined in the current study. If the
influence of affective information on social cognitive processes like trustworthiness appraisal is amenable to change, this change itself may then act as a mechanism for other treatments gains (e.g., improved interpersonal stability). Incorporating behavioral tasks and consideration of social cognition into treatment research may provide valuable insights into how treatment works and how it might work better.

**Conclusion**

Interpersonal and affective instability are two core characteristics of BPD. Vociferous debate regarding which of these instabilities lies at the core of BPD wages on and may not ever reach resolution. Rather than tackle this debate directly, the current work sought to combine the study of interpersonal dysfunction and affective instability to better understand how these instabilities may interact. More specifically, I sought to examine the influence of affective instability on a social cognitive process, and in turn to examine the influence of these processes on interpersonal functioning.

In this dissertation I have presented findings from two behavioral tasks designed to further elucidate trustworthiness processing biases in BPD. The current dissertation confirms the idea that individuals with borderline personality pathology have a bias to rate others as less trustworthy. Additionally, this work shows that negative affective information exerts exaggerated influence on these ratings for individuals with high borderline features, compared to individuals with low borderline features. Importantly, and contrary to study predictions, trustworthiness bias and susceptibility to priming effects from negative affective information were largely not related to measures of social functioning for individuals with borderline features, or for participants overall regardless of borderline features. These findings call into question the relevance of social cognitive biases and negative affective interference for our understanding of how individuals
with borderline pathology function in their social environments. Regardless, the current findings add meaningfully to a growing wealth of literature showing that negative affective information interferes with a variety of cognitive processes for individuals with borderline pathology. As such, individuals with borderline pathology are likely to have more difficulty regulating cognitive processes in the context of distress.
References


Bell-Pringle, V.J., Pate, J.L., & Brown, R.C. (1997). Assessment of borderline personality disorder using the MMPI-2 and the Personality Assessment Inventory. *Assessment, 4*(2), 131-139.


dysfunction, and borderline personality disorder. *The Journal of Nervous and Mental Disease, 197*(11), 816-821.


Mikolajczak, M., Pinon, N., Lane, A., de Timary, P., & Luminet, O. (2010). Oxytocin not only increases trust when money is at stake, but also when confidential information is in the balance. *Biological Psychiatry, 85*, 182-184.


personality disorder subjects with a history of sustained childhood abuse. *Biological Psychiatry, 52*(11), 1102-1112.


Wingenfeld, K., Menschach, C., Rallkoetter, N., Schosser, N., Schafrath, C., Beblo, T., & Dienstbier, K. (2009). Relationship between coping with negative life events and...


Appendix A

General Trust and Trustworthiness Items

### Trust Items

<table>
<thead>
<tr>
<th></th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am generally a trusting person</td>
</tr>
<tr>
<td>2</td>
<td>I give people the benefit of the doubt in deciding whether to trust them</td>
</tr>
<tr>
<td>3</td>
<td>I lose trust for others easily</td>
</tr>
<tr>
<td>4</td>
<td>Once someone loses my trust, it is very difficult for me to trust them again</td>
</tr>
<tr>
<td>5</td>
<td>I am too trusting of others</td>
</tr>
<tr>
<td>6</td>
<td>I do not trust anyone until they have done something to earn that trust</td>
</tr>
<tr>
<td>7</td>
<td>I am on the lookout for signs that I shouldn't trust people</td>
</tr>
</tbody>
</table>

### Trustworthiness Items

<table>
<thead>
<tr>
<th></th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My friends can trust me with their secrets</td>
</tr>
<tr>
<td>2</td>
<td>My family can trust me with their secrets</td>
</tr>
<tr>
<td>3</td>
<td>When someone lends me money, I pay it back</td>
</tr>
<tr>
<td>4</td>
<td>I am a trustworthy person</td>
</tr>
<tr>
<td>5</td>
<td>People can trust me to tell the truth</td>
</tr>
<tr>
<td>6</td>
<td>People can trust me not to cheat</td>
</tr>
<tr>
<td>7</td>
<td>I am a liar</td>
</tr>
</tbody>
</table>
Appendix B

McLean Assessment of Rejection Sensitivity

(circle the number that best characterizes you)

<table>
<thead>
<tr>
<th></th>
<th>Like</th>
<th>(Much)</th>
<th>(Some)</th>
<th>Unlike</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

   1. I keep my antenna up for signs of rejection

   2. I would not compromise my standards to gain others’ approval

   3. I often need reassurance that I haven’t offended others

   4. Sometimes I perceive rejection too readily

   5. I’m OK with the idea that some people won’t like me

   6. Criticisms often make me feel like someone dislikes me

   7. Being fair is more important to me than being popular

   8. I try to avoid situations where I might be rejected

   9. Separations often feel like rejections

  10. It doesn’t bother me that much to have someone angry at me
Appendix C

Social Interaction Initiation Scale (SIIS)

Social Interactions Initiated by the Participant:

Please consider the following activities in terms of how often you engage in them during a typical week. Indicate the typical number of times per week you do each activity, and then whether you would prefer to do this activity the same amount, less often, or more often.

Frequency: 0, 1, 2, 3, 4, 5+

Positive:

1. Initiated plans with friends such as dinner, going out, or attending a party.
2. Spoken with a family member.
3. Had a brief conversation or small talk with a stranger.
4. Initiated spontaneous plans with someone (lunch, coffee, walk, etc).
5. Successfully resolved an argument with someone you are close to.
6. Spoken to a friend via phone or email
7. Given a compliment
8. Spoken with a friend in person
9. Made a new friend
10. Flirted with someone, or asked someone out on a date?
11. Spent time with other people in a group setting

Negative:

1. Had an argument with someone you are close to
2. Given an ultimatum or a threat to someone else
3. Had an argument with a store clerk or other serviceperson
4. Initiated an argument with someone at your place of work or school
5. Honked your horn at someone else in traffic, not out of concern for safety.
6. Wished that you were spending time alone
7. Kept someone waiting more than 10 minutes
8. Kept someone waiting more than 30 minutes
9. Did not attend a social gathering to which you were invited and could have attended.
10. Spent time with someone with whom you have previously had a serious disagreement
11. Were unsatisfied with the quality of your social interactions.
12. Walked away from a social interaction feeling bad or awkward about it.
13. Ended a friendship
14. Regretted the way that you behaved in a social situation
15. Did something to get back at someone
16. Lied to get out of a social commitment
17. Made up an excuse to miss work
18. Was mean to someone for no reason
19. Criticized someone unnecessarily

**Leisure Satisfaction:**

1. I was satisfied with the quality of my social leisure time.

Would you generally agree with this statement? (responses on 5-point scale)

**Social Interactions Initiated by Others:**

Please consider the following activities in terms of how often you engage in them during a typical week. Indicate the typical number of times per week you do each activity with a member of your peer group and then whether you would like others to initiate these activities the same amount, less often, or more often.

Frequency: 0, 1, 2, 3, 4, 5+

**Positive:**

1. Been invited to a planned activity with a group
2. Been called by a family member
3. Responded to a stranger's initiation of conversation or small talk
4. Been invited to spontaneous plans with someone (coffee, lunch, etc.)
5. Been called or contacted by a friend to "catch up"
6. Been asked on a date, or was flirted with.
7. Received socially related emails or text messages

**Negative:**

1. Been given an ultimatum or other threat
2. Not been invited to a social gathering which your friends had organized
3. Felt someone you did not know was being rude to you
4. Had someone cancel plans you initiated
5. Felt disappointed by your social opportunities
6. Felt overwhelmed by social experiences
7. Felt that a sales clerk or service person was rude or unkind to you
8. Felt that a person or group of people are “out to get you”
9. Were unsatisfied by the number of friends that you have
10. Were upset by the way in which others behaved during social leisure time.
Appendix D

Nobody's Perfect Questionnaire

Nobody is perfect, and many of us have done some of the following things. Please indicate how many times you have done each of the following things.

Circle the number that corresponds to your answer.

The experimenter will not look at this questionnaire until after the study is complete, so please make sure you fill out every question.

You will not be judged based on your answers here, so please answer as truthfully as possible.

<table>
<thead>
<tr>
<th></th>
<th>I have never done this.</th>
<th>I did this once.</th>
<th>I have done this more than once.</th>
<th>I do this fairly frequently, or I have done this more than a few times.</th>
<th>I do this all the time, or I have done this many times.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Telling someone they look better than they actually do</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2 Driving 5 miles/hour over the speed limit</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3 Shoplifting</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4 Driving very aggressively or well over the speed limit</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5 Lying to get out of doing something I don't want to do</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6 Calling in to work sick when I'm not sick</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7 Cheating on a partner/spouse/boyfriend/girlfriend</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8 Telling someone I'm more accomplished than I really am</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9 Cheating on a test</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10 Saying something mean to hurt someone else</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11 Lying on my taxes</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12 Giving false information on a job application</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13 Passing by someone who needs help that I can give</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14 Stealing something from a person I know</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15 Never returning a library book</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16 Not paying a parking ticket</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17 Setting it up so someone else takes the fall for my mistake</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18 Fighting with someone else physically</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19 Losing my temper with a child and yelling or screaming</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20 Losing my temper with another adult and yelling or screaming</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>21 Passing my work off on someone else</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>22 Sabotaging a friend</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>23 Cutting in line</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>24 Ripping someone off or taking advantage of someone's weakness</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>25 Not giving up a seat for someone elderly or disabled</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>26 Not holding a door for someone behind you</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>27 Littering</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>28 Pusposely avoiding someone so I don't have to talk to them</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Appendix E

**Being Nice Questionnaire**

Nobody is all bad, and we have all done things that are good or that benefit others. Please indicate how many times you have done these things. Circle the number that corresponds to your answer.

You will not be judged based on your answers here, so please answer as truthfully as possible.

<table>
<thead>
<tr>
<th></th>
<th>I have never done this.</th>
<th>I did this once.</th>
<th>I have done this more than once.</th>
<th>I do this fairly frequently, or I have done this more than a few times</th>
<th>I do this all the time, or I have done this many times.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Taking on extra work to help a coworker</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Lending money to a friend in need</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Helping an elderly person you don't know</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Volunteering to help others</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Donating money to charity</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Holding a door open for someone else</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>Sending flowers to a sick friend or following a funeral</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>Caring for another person when they are sick</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>Giving a surprise gift to someone else</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Calling someone else to see how they are doing</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>Taking the blame for someone else to help them</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>Doing a last minute favor for a friend</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>Picking up trash that someone else left behind</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>Doing your part to care for the environment</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>Giving emotional support to a friend during a challenging time</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>Making someone feel special</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>Giving up my place in line to help someone in a rush</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>Giving my seat up for someone elderly or disabled</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>Admitting to my mistakes</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>Defending someone else</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>21</td>
<td>Stopping at crosswalks to let pedestrians cross</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>