Towards a comprehensive understanding of Perceived Criticism and its predictive validity: Three studies investigating the correlates and consequences of PC

Permanent link

Terms of Use

Share Your Story

Towards a comprehensive understanding of Perceived Criticism and its predictive validity:

Three studies investigating the correlates and consequences of PC

A dissertation presented

by

Steven Arthur Melbourne Felix

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

September 2018

Committee Members

Jill M. Hooley, D.Phil. (Advisor, Harvard University)

Christine I. Hooker, Ph.D. (Rush University)

Matthew K. Nock, Ph.D. (Harvard University)

Richard J. McNally, Ph.D. (Harvard University)

Towards a comprehensive understanding of Perceived Criticism and its predictive validity:

Three studies investigating the correlates and consequences of PC

Abstract

Perceived criticism (PC) is a robust predictor of poor clinical outcomes for a range of mental disorders, and it predicts these outcomes better than objective measures of criticism. Though PC has the potential to be a valuable clinical tool, little is known about why PC predicts relapse, or other unfavorable transdiagnostic outcomes. The present dissertation begins to address this question by investigating the informational content of PC. Specifically, a series of studies examines the extent to which PC provides information about relationship experiences as opposed to traits or characteristics of the perceiver. In paper 1, we find evidence that PC provides both types of information; specifically, we find that people perceive similar levels of criticism across a number of important relationships. Furthermore, PC in one relationship predicts PC in other relationships just as strongly as relationship-specific characteristics, such as relationship satisfaction. In paper 2, we investigate whether global PC ratings are related to well-being because they index daily experiences of criticism exposure. We find that global PC is only modestly related to daily criticism and that global PC incrementally predicts daily feelings of social support and connectedness after controlling for depression, relationship satisfaction, and daily perceptions of criticism. These findings suggest that PC may predict some aspects of well-being due to shared variance with depressive symptomatology, and also that PC itself may be a unique indicator for the ways people experience and evaluate their social worlds each day. In paper 3, we test the hypothesis that people reporting high PC would demonstrate enhanced recall for negative information when it is processed self-referentially. This hypothesis was not supported, suggesting that PC cannot be dismissed merely as a negative bias in the way a person recalls information about the relationship. Together, these studies provide a unique perspective on the meaning of the PC construct. These studies indicate that, to a significant

extent, PC reflects clinically-relevant processes and characteristics at the level of the individual. These processes may be primarily related to ways of perceiving and interpreting *social* experiences. Thus, potential interventions for PC must incorporate both individual and family components.

Table of Contents

Acknowledgements

This work could not have been completed without significant support, generosity, and brilliance of several individuals and groups. My first advisor, Dr. Christine Hooker, inspired me to think big and pursue novel, ambitious ideas. She helped me find optimism and enthusiasm when I had trouble seeing the value of my research. Both Dr. Hooker and my second advisor, Dr. Jill Hooley, also supported me when I wanted to take a leave of absence to pursue other interests; and they welcomed me back with open arms when I wanted to return. I'd also like to thank Dr. Hooley for taking me under her wing when an amazing new opportunity at Rush University beckoned Dr. Hooker to Chicago. Dr. Hooley challenged me to be intentional about every idea, design, and measure included in my research, and provided a platform for me to think deeply and rigorously about the fascinating topics contained in this dissertation.

I would also like to thank all members of my dissertation committee, including Drs. Hooker, Hooley, and Nock, for providing extremely helpful feedback on this research, as well as Dr. Rich McNally, for serving as an outside examiner.

I am also extremely grateful to members of the Clinical Research Laboratory, including Dr. Sara Masland, Dr. Ken Allen, Kathryn Fox, Tina Chou, Chelsea Boccagno, and Shirley Wang, as well as members of the Social Neuroscience and Psychopathology Lab, particularly David Dodell-Feder, Cora Mukerji, and Erik Nook. Each of them helped me along my path by providing feedback, advice, and camaraderie. I am also very thankful for my talented research assistants, Matthew Abrams, Matthew Yung, Ruichen Sun, Summer Peterson, Gita Bhattacharya, and Kendall Baldwin, without whom I would not have been able to conduct this research.

I also thank the organizations that funded my research, including Sackler Foundation, the Department of Psychology at Harvard University, the American Psychological Association of Graduate Students, and the Psi Chi National Honor Society in Psychology.

Finally, I thank my family: my wife, Heather, who gave me endless support and patience throughout graduate school; my children, Riley and April, who enriched my life with laughter, happiness, and perspective; my mother and father in-law, who provided life-saving babysitting for me to get work done and keep my sanity; and my parents, who supported my educational journey since childhood.

Introduction

The quality of our relationships, particularly those with our intimate partners, plays a significant role in our well-being. Compared to those in more negative or ambivalent relationships, individuals in positive, satisfying relationships demonstrate better health across a variety of mental and physical domains, such as better cardiovascular and immune functioning, reduced depression and anxiety, fewer chronic health issues, lower likelihood of being disabled, lower mortality, and reduced perceptions of pain (for review, see Kiecolt-Glaser & Newton, 2001). These associations are not merely correlational; both positive and negative dimensions of relationship functioning exert these effects in direct and specific pathways (Fincham & Rogge, 2010; Kiecolt-Glaser & Newton, 2001). Positive aspects of relationship functioning include elements such as supportive behaviors (Feeney & Collins, 2014), equitable contributions to the relationship (Gleason, Iida, Shrout, & Bolger, 2008), responsiveness (Reis, Clark, & Holmes, 2004), and intimacy (Debrot, Cook, Perrez, & Horn, 2012). Although positive aspects of relationship functioning undoubtedly contribute to health and may provide buffering functions in times of stress, the negative aspects of relationships more strongly and consistently predict health and relationship outcomes (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Seidman, 2012). Negative dimensions of relationships include behaviors such as hindrance and undermining (Rafaeli, Cranford, Green, Shrout, & Bolger, 2008), hostility (Fincham & Rogge, 2010), blaming (Bradbury & Fincham, 1990), and criticism (Hooley, 2007).

Criticism is particularly harmful to mental health. A seminal study by G. W. Brown, Birley, and Wing (1972) investigated why patients with schizophrenia were more likely to relapse into florid psychosis after being discharged to live with spouses, siblings, or parents. Based on previous findings, the researchers hypothesized and confirmed that family environments characterized by high amounts of criticism, hostility, and emotional involvement—together referred to as Expressed Emotion (EE)—increased the risk of relapse independently from other factors such as illness severity, duration of illness, or symptomatology. Although it was initially

thought that EE was a risk factor specifically for relapse in schizophrenia, subsequent investigations have demonstrated that EE is a robust predictor of relapse for many mental disorders, including schizophrenia, depression, eating disorders, and substance use disorders (Brown et al., 1972; Butzlaff & Hooley, 1998; for review see Hooley, 2007). EE also predicts poor treatment outcomes in posttraumatic stress disorder (Tarrier, Sommerfield, & Pilgrim, 1999).

The ability of EE to predict such a wide array of important clinical outcomes suggests that it could be used in clinical settings to identify patients with elevated risk for such outcomes. Unfortunately, EE is not easily assessed. EE is measured using the Camberwell Family Interview (CFI; Leff & Vaughn, 1985). The CFI is a semi-structured interview with a patient's relative. The interview is recorded and subsequently coded by trained raters for its three components. Criticism is assessed by counting the number of critical comments about the target person; hostility is rated as being present or not; and emotional overinvolvement is rated on a 0-5 scale. Though the CFI has many strengths—including a high degree of reliability and objectivity—this comes at a cost. With regards to training, it takes 2 to 4 weeks to become a reliable CFI coder. It can also be difficult to find training, as there are few CFI instructors. Individuals interested in learning the CFI may have to travel internationally to find a CFI instructor. Administration and scoring of the CFI also represent a significant burden to both patients' relatives and the assessor: It takes 1 to 2 hours to administer the CFI and many more hours to review the audio recordings for scoring purposes. Together, these requirements of the CFI make it impractical for routine clinical use.

In an attempt to simplify the CFI rating process, Hooley and Richters (1991) investigated whether the CFI would retain its predictive and concurrent validity if undergraduates, as opposed to trained raters, made ratings of EE for spouses of depressed patients. They found that although undergraduates' ratings of EE correlated with those of trained raters, undergraduates' ratings did not predict 9-month relapse. Hooley and Richters (1991)

2

emphasized the importance that new measures of EE must be developed with the goal of improving prediction of relapse, not of finding an easier or better measure of EE.

Hooley and Teasdale (1989) developed the Perceived Criticism (PC) measure in a similar attempt to simplify the process of measuring EE. PC is a subjective, global self-report evaluation of the amount of criticism a person perceives in a particular relationship. It is measured by having a respondent identify his or her most important relationship and then rate how critical that person is on a Likert scale from 1 ("not at all") to 10 ("very critical"). Compared to using the CFI, measuring PC requires no training and takes only a few seconds to complete. This makes it much easier to use in both clinical and research settings. Hooley and Teasdale (1989) focused on criticism, as opposed to hostility or emotional overinvolvement, due to the consistent finding that criticism was the most important component of EE in predicting patients' clinical outcomes.

The ultimate test of PC is not its ease of use but its ability predict mental health outcomes compared to objective measures of EE (Hooley & Richters, 1991). Testing PC against EE in a sample of depressed patients, Hooley and Teasdale (1989) found that PC was a better predictor of relapse than EE. All patients reporting PC of 6 or higher relapsed. In contrast, none of the patients reporting 2 or lower relapsed. Subsequent research has replicated and extended these findings by demonstrating that PC predicts relapse, treatment outcomes, and symptom increases across a variety of mental disorders over and above EE or other global relationship evaluations, such as relationship satisfaction (for review, see Masland & Hooley, 2015). These findings provide the main impetus for continued use of PC as a marker for vulnerability to relapse and other clinical outcomes.

However, there are a number of unanswered questions related to PC that prevent it from being truly useful in clinical settings. Primary among these questions is how and why PC predicts these many negative outcomes. Does PC play a causal role in the exacerbation of mental health problems, perhaps through stress pathways? Or is PC merely a byproduct of

other vulnerabilities or pathological processes? Only by understanding the mechanisms connecting PC to clinical outcomes can effective interventions can be designed and implemented. However, before questions of mechanisms can be addressed, a more fundamental question must be answered: What is PC truly measuring? What *is* the PC construct? One of the first steps in understanding and harnessing the predictive power of PC involves establishing its construct validity. Does PC reflect genuine relationship criticism? Does it measure general relationship distress? Or does PC index characteristics of the individual, such as cognitive or interpretative biases? Although PC is clearly face-valid, the question of what exactly is being measured via a PC rating is less straightforward. Certainly, research has begun to address this question (Chambless & Blake, 2009; Masland, Drabu, & Hooley, *in press*; Riso, Klein, Anderson, Ouimette, & Lizardi, 1996). Nonetheless more questions than answers remain. The purpose of this dissertation is to create a more holistic understanding of the PC construct and its informational content, with an eye towards understanding why PC predicts poor clinical outcomes.

Validity of the PC Construct

An important test of PC's construct validity is its convergence with other measures of criticism. Investigations into the convergent validity of PC have yielded mixed results. To understand these findings, it is helpful to distinguish standardized and non-standardized measures of criticism. The most common standardized measures of criticism and hostility include the CFI, the Interaction Rating System (KPI; Hahlweg & Conrad, 1983), and the Five Minute Speech Sample (FMSS; Magaña et al., 1986). In their seminal study on PC among patients with depression, Hooley and Teasdale (1989) found a significant correlation of $r = 0.51$ between PC and EE, as measured by the CFI. However, correlations between PC and the individual criticism and hostility components of the CFI were not significant. In contrast, Lee, Barrowclough, and Lobban (2014) did find significant correlations between PC and CFI-rated criticism, hostility, and overall EE in a sample of patients experiencing first-episode psychosis.

4

With respect to CFI-rated criticism, these conflicting findings may not merely be a difference of sample sizes. Though Lee and colleagues included a larger sample size (N=60 vs. N=23), the effect size was much larger. Using multi-level models, Gerlsma, van Duijn, Hale, and van Hout (2009) also found that the number of critical comments from a spouse during the FMSS was significantly related to PC even after controlling for the perceiver's own critical comments, depressed mood, and marital satisfaction. In a second study by these authors, KPI-rated criticism was found to be related to PC, but this relationship was no longer significant after controlling for the perceiver's own nonverbal negative behavior towards his or her spouse.

Some studies have failed to demonstrate a relationship between PC and standardized measures of EE. Using structural equation modeling, Chambless, Bryan, Aiken, Steketee, and Hooley (2001) found no relationship between PC and CFI-rated hostility in a sample of outpatients suffering from obsessive-compulsive disorder or panic disorder with agoraphobia. Similarly, Chambless and Blake (2009) found no relationships between PC and CFI-rated criticism, hostility, or emotional overinvolvement and KPI-rated criticism. Furthermore, some studies using the FMSS failed to detect a relationship between PC and EE status (i.e., high vs. low) (Bachmann, Bottmer, Jacob, & Schröder, 2006) or between PC and the number of critical comments (von Polier et al., 2014).

The inconsistent relationship between PC and standardized measures of EE might result from a number of possibilities, many of which do not preclude its construct validity. For instance, the CFI, KPI, and FMSS allow observers to see only a brief, singular perspective on the relationship that exists between two people. In contrast, PC allows a person to draw upon many years of experiences and interactions with his/her relative when providing a PC rating. Another possibility has to do with measurement itself: Objective measures typically rely on frequency counts of the number of critical comments made. When counting these critical comments, standardized definitions of criticism must be applied to all interviewees. Thus, these approaches cannot take relationship-specific nuances into account. PC, in contrast, reflects an overall

5

impression or opinion about a relative's criticism that has formed over many years. Furthermore, this impression allows each person to use their own definition of criticism and their own interpretation of their partners' behavior, for instance to taking intentions into consideration. In general, these explanations for PC's predictive validity imply that although PC may be different from what is measured in standardized assessments of criticism, it may still be a valid measure of the *experience* of criticism.

Supporting this idea, a number of studies have demonstrated that PC converges with *subjective* or non-standardized ratings of criticism from untrained raters. For instance, PC is correlated with relatives' (e.g., spouses') own perceptions of how critical or hostile they are towards the patient, both globally and in laboratory interactions (Chambless & Blake, 2009; Chambless et al., 2001; Smith & Peterson, 2008). PC also converges with untrained observers' overall impressions of how much *destructive criticism* they perceive in a spousal interaction (Chambless & Blake, 2009). Similarly, Smith and Peterson (2008) found a small to medium correlation between global PC and observer-rated criticism in a marital interaction. This convergence with untrained raters' perceptions of criticism suggests that PC is picking up on true relationship criticism and thus cannot be dismissed out-of-hand as only reflecting biased perceptions or interpretations. Furthermore the contrast in associations between standardized and unstandardized assessments of criticism suggests that PC may reflect global *impressions* of criticality, rather than frequency or volume of criticism. Hooley and Teasdale (1989) suggested that PC reflects how much criticism is "getting through" (p. 234). From this perspective, it is not surprising that standardized methods of counting critical comments (e.g., the CFI, KPI, FMSS)—which do not take into consideration the factors that determine whether a behavior or comment is subjectively perceived as critical and which rely on only a thin slice of a relative's behavior—do a poor job of assessing an individual's subjective experience of criticism in a relationship.

Present Dissertation and Research Questions

One of the central tasks facing researchers interested in PC, as well as clinicians interested in reducing risk of relapse in their high-PC patients, is to understand the meaning of a PC rating. Though several studies have addressed this question, relatively few studies have sought to develop a holistic conceptualization of PC. Instead, they have generally focused on testing PC's convergent validity with objective measures criticism. As described above, findings from these studies indicate that objective impressions of a relative's criticality may account for, at most, 20 – 25% of the variance in PC. In other words, much is left to be explained.

The purpose of the present dissertation is to define more precisely how we should conceptualize the PC construct. A key component of this endeavor is to understand how PC is distinct from objective criticism; that is, what can a PC rating tell us about a person and his or her risk for adverse outcomes that objective criticism cannot? Addressing this question will not only provide a more holistic understanding of the PC construct but also highlight potential mechanisms explaining PC's prediction of transdiagnostic outcomes. In Paper 1, we build upon this Introduction by providing a comprehensive summary of the current state of knowledge around the PC construct, including known correlates and possible explanations for their relationships. The research described in Paper 1 expands upon this knowledge by pitting relationship-level variables directly against person-level variables in statistical models in order to test whether PC should be thought of as a relationship or trait variable. Paper 2 assesses the relationship between traditional PC ratings ("global" PC) and *daily* perceptions of criticism in order to understand the extent to which PC predicts well-being simply because it indexes daily experiences of criticism. Specifically, Paper 2 assesses whether the relationship between global PC and daily well-being can be explained by daily impressions of partner criticism. Lastly, Paper 3 investigates whether high PC is characterized by biased recall for negative information, as this may not only explain how PC ratings are formed but also why they predict clinical outcomes.

In summary, I present the following three research questions and studies:

Question 1: Do people perceive similar amounts of criticism in all of their important relationships? Are PC ratings better explained by relationship variables or individual variables?

Paper 1. Felix, S. A. M., Masland, S. R., & Hooley, J. M. (2017). Is perceived criticism more than a relationship variable?. *Manuscript under revision.*

Question 2: Is the relationship between PC ratings and daily well-being explained by daily experiences of criticism?

Paper 2. Felix, S. A. M., Hooley, J. M., & Hooker, C. I. (2017). Global perceived criticism ratings incrementally predict daily social appraisals over daily criticism exposure. *Manuscript under revision.*

Question 3: Is PC associated with biased recall for negative semantic information?

Paper 3: Felix, S. A. M., & Hooley, J. M. (2018). Perceived criticism is not merely a negative memory bias. *Manuscript in preparation.*

Paper 1: Is perceived criticism more than a relationship variable?

Steven A. M. Felix, Sara R. Masland, & Jill M. Hooley

Manuscript under revision

Abstract

Background: A reliable predictor of poor clinical outcomes, perceived criticism (PC) has traditionally been conceptualized as a relationship-specific construct. However, this conceptualization does not fully account for PC's prediction of clinical outcomes. The present study examines the extent to which PC characterizes the person providing the rating.

Design: Amazon MTurk workers (n=197) who were in romantic relationships first provided PC ratings for 5 important relationships and then completed a battery of self-report measures assessing both individual and relationship variables.

Results: PC was related to both individual and relationship variables. People also reported similar levels of PC across relationships such that average PC from a set of relationships (e.g., mother, father, friend) was one of the strongest predictors of PC another relationship (e.g., partner). Individual variables such as dysfunctional attitudes significantly predicted average PC.

Conclusions: PC is not merely a measure of criticism in a single relationship. High PC ratings indicate an elevated risk profile, including greater cognitive, affective, and social vulnerabilities.

Introduction

Perceived Criticism (PC) is a clinical construct that indexes the amount of criticism a person perceives in his or her closest relationship. It is measured using the single item Perceived Criticism Measure (PCM; for review see Masland & Hooley, 2015). Respondents first identify their most important relationship and then rate how critical that person is using a 10-point Likert scale. This simple self-report measure reliably predicts poor outcomes across several clinical disorders, including 6 and 12-month relapse in first-episode psychosis (Lee et al., 2014); less improvement in target problems (Chambless & Steketee, 1999) and greater post-treatment symptom severity (Renshaw, Chambless, & Steketee, 2003) in obsessive compulsive disorder and agoraphobia; higher rates of relapse and shorter time to relapse in substance use disorders (Fals-Stewart, O'Farrell, & Hooley, 2001); likelihood of hospital admission in bipolar disorder (Scott, Colom, Pope, Reinares, & Vieta, 2012), and relapse in depression (Hooley & Teasdale, 1989; Kwon, Lee, Lee, & Bifulco, 2006). Typically, the PCM also predicts clinical outcomes incrementally over symptomatology, relationship satisfaction, and 'objective' measures of criticism, such as the frequency count of critical comments made by a patient's relative during the Camberwell Family Interview (CFI; Leff & Vaughn, 1985) (Hooley, 2007). The reason for PC's incremental predictive validity is not yet clear. This is partially due to lingering questions about exactly what PC is measuring. It is generally assumed that PC is a relationship-specific construct. However, its incremental predictive validity suggests PC ratings may also be tapping other clinically-relevant constructs. In the following sections, we examine three general types of information that PC ratings might provide and discuss how this information may add to our understanding of the link between PC and clinical outcomes.

PC as a Measure of Objective Criticism

The original intent of the PCM was to provide a short-cut measure of criticism (Hooley & Parker, 2006). Criticism is the key element of expressed emotion (EE), a set of characteristics of the family environment that has been reliably linked to relapse for disorders such as

schizophrenia and depression (Butzlaff & Hooley, 1998; Hooley, 2007). Assessing EE in the

conventional manner requires administering the CFI to each patient's key relative. The CFI,

which takes approximately 90 minutes, must then be coded by a trained-rater in order to derive

an independent assessment of criticism. Hooley and Teasdale (1989) first developed the PCM

to provide a more rapid assessment of criticism from the patient's perspective. They reported

that the PCM predicted adverse clinical outcomes even better than ratings of EE derived from

the CFI. The significant correlation between patients' subjective ratings of criticism from their

relatives and EE ratings obtained from the CFI further suggested that PC ratings were indeed

measuring important aspects of the family environment. However, there was only a modest

correlation ($r =.27$) between PC and the objective CFI rating of criticism. Subsequent studies

have reported mixed-findings with regards to PC's relationship with objective measures of

criticism. Some studies report significant correlations between PC and standardized measures

of criticism and hostility (Gerlsma et al., 2009; Lee et al., 2014), although this is not invariably

the case (Bachmann, Bottmer, Jacob, & Schröder, 2006; Chambless & Blake, 2009;

Chambless, Bryan, Aiken, Steketee, & Hooley, 2001; von Polier et al., 2014).

Why might PC ratings effectively predict clinical outcomes despite having only weak or

modest correlations with objective criticism? One possibility is that PC provides a more accurate

reflection of a person's *experience* of criticism. Objective measures of criticism (such as the one

derived from the CFI) are based on relatively brief interviews or observational periods. These

may not be sufficient to provide a thorough or representative sampling of a dyad's general

interactional patterns (Renshaw, 2008). A PC rating, in contrast, is often based on years of

interacting and living with a relative or spouse; as such it may be more reliable. PC ratings may

also be more accurate because they allow each person to use his or her own definition of

criticism and own interpretations of a relative's behavior (e.g., by taking partners' intentions into

account). Standardized measures must rely on fixed definitions of criticism and so may be

unable to detect relationship-specific nuances or differences in how individuals evaluate whether a comment is (or is not) critical.

In contrast to the inconsistent findings with regards to objective or standardized measures of criticism, PC is consistently correlated with more subjective measures of criticism. For instance, PC correlates with spouses' own perceptions of how critical or hostile they are towards their partners (Chambless & Blake, 2009; Chambless et al., 2001; Smith & Peterson, 2008). PC also correlates with untrained raters' overall impressions of destructive criticism in couples' laboratory interactions (Chambless & Blake, 2009). PC may correlate with these subjective measures more strongly and consistently because, like PC, they reflect more global and intuitive *impressions* of criticism rather than a strict count of frequency. Importantly, the consistency with which PC correlates with these subjective measures demonstrates convergent validity and provides some support for PC as a measure of true relationship criticism. Nonetheless, the largest of these correlations (r = .54, Smith & Peterson, 2008) indicates that genuine criticism and hostility—whether measured objectively or subjectively—may account for at most 20-30% of the variance in PC ratings. The large amount of unexplained variance in PC, even when considering the upper bound of its relationship with other measures of criticism, suggests that PC may be more than a measure of criticism.

PC as a Measure of General Relationship Quality

PC may also index the overall health and quality of a relationship. PC ratings for romantic partners are robustly associated with relationship adjustment or satisfaction. Studies have consistently found correlations between the PCM and the Dyadic Adjustment Scale (DAS; Spanier, 1976) ranging between -0.49 and -0.54 (Chambless & Blake, 2009; Hooley & Teasdale, 1989; Smith & Peterson, 2008). Objective measures of criticism, in contrast, are not reliably correlated with relationship adjustment (Chambless & Blake, 2009; Hooley & Teasdale, 1989; Smith & Peterson, 2008). Thus, PC's predictive validity may partially be due to its assessment of a person's overall positive or negative experience in a relationship. Even so, it is

13

important to keep in mind that PC shares only about 25% of its variance with relationship

adjustment (based on its largest correlations with DAS). PC also continues to predict clinical

outcomes even after relationship adjustment or satisfaction is statistically controlled (Hooley &

Teasdale, 1989; Renshaw, 2008). Similarly, factors such as relationship attributions account for

a significant amount of variance in PC ratings even after relationship adjustment are statically

controlled (Peterson & Smith, 2011). Thus, although PC ratings may be influenced by

relationship adjustment, it is not merely a measure of global relationship factors like adjustment

or satisfaction. Other explanations are needed to understand why PC provides incremental

predictive validity.

PC as a Measure of Individual Traits

Theoretical accounts of PC have generally assumed that it is a relationship-specific

construct, and the prior two sections highlight reliable evidence to support this assumption.

However, this may not be the whole story, as PC has been associated with individual traits of

raters themselves. For instance, PC has been linked with cognitive biases: Individuals reporting

high PC make more negative interpretations of ambiguous information and demonstrate

difficulty shifting attention away from negative social stimuli (angry faces) (Masland, Hooley,

Tully, Dearing, & Gotlib, 2015). Importantly, these biases may confer vulnerability to

psychopathology: Both attentional and interpretative biases have been documented in

depressed and remitted-depressed samples, as well as in samples at high risk for depression,

such as children whose mothers are depressed (Gotlib & Joormann, 2010; Jacobs, Reinecke,

Gollan, & Kane, 2008; Joormann, Talbot, & Gotlib, 2007; Joormann, Yoon, & Zetsche, 2007;

Williams, Watts, MacLeod, & Mathews, 1997). Thus, people reporting high PC may be at risk for

poor clinical outcomes not only because of a harmful relationship, but also because they have

these more general cognitive vulnerabilities.

PC is also associated with aberrations at the neural level. Hooley, Siegle, and Gruber

(2012) found that women who reported high versus low PC demonstrated reduced activity in the

dorsolateral prefrontal cortex (DLPFC) and greater activity in the amygdala while listening to criticism (but not praise) from their mothers. The LPFC has been associated with cognitive control of emotional information, specifically down regulation of neural response to negative stimuli (Hooker, Gyurak, Verosky, Miyakawa, & Ayduk, 2010; Ochsner & Gross, 2005), and the amygdala is associated with ascribing emotional value to information (Morrison & Salzman, 2010). These aberrations may have important clinical consequences: Reduced LPFC activity during laboratory stressors has been associated with greater rumination and negative mood after real-world conflict (Hooker et al., 2010). Studies have also found similar aberrant DLPFC and amygdala activity among fully remitted-depressed patients (Hooley et al., 2009; Hooley, Gruber, Scott, Hiller, & Yurgelun-Todd, 2005), and PC continues to predict this neural activity even after controlling for depression (Hooley et al., 2012). Though the temporal and causal relationships between PC and these aberrations are unknown, it is possible that these neural abnormalities could make individuals less resilient in the face of criticism or other stressful experiences.

PC has also been linked with criticality bias, the tendency to over-perceive criticism when none was intended (Peterson, Smith, & Windle, 2009; Smith & Peterson, 2008). Smith and Peterson (2008) recruited married couples to engage in a conversation about what they would change about their partners. The interactions were recorded and subsequently rated by both trained observers and each of the partners for the amount of criticism observed, intended, or perceived (depending on who was doing the rating). Bias was calculated as the difference between perceptions of being criticized and either objective raters' observations of criticism or partners' intended criticism. The authors found that participants who tended to over-perceive criticism also reported higher global PC, suggesting that PC is to some extent rooted in a cognitive bias. However, the authors noted that PC continued to predict dysphoria even after controlling for criticality bias. The authors suggest this may be because PC acts as a measure of both individual bias and objective criticism. Similar to PC, it is unclear whether criticality is

15

trait-like or relationship-specific. That is, do people demonstrate similar levels of criticality bias across different relationships? If criticality bias tends to be trait like, it could cause PC to appear more trait like as well.

Indeed, existing evidence suggests that PC may represent an individual difference trait. One study found significant correlations among undergraduates' PC ratings for 5 meaningful relationships (White, Strong, & Chambless, 1998). Students tended to perceive similar amounts of criticism from parents (r = .42), and from peers (α = .75 among 3 peers). Similarities in PC ratings across relationship have implications for the PC construct as well as its predictive validity. First, they suggest that individual-level processes (e.g., biases) play a substantive role in how people make PC judgments in various relationships. Identifying the processes accounting for similarity in PC ratings is crucial, as these traits may hold the key to understanding why PC predicts clinical outcomes. These processes may represent a latent vulnerability to psychopathology, of which PC is one result. Second, they suggest that people have similar experiences in their relationships. Thus, PC may indicate risk of adverse mental health outcomes because it describes the overall quality of many social relationships.

Though PC has been linked with some cognitive and behavioral biases, several studies have failed to find correlations between PC and a variety of other clinically-relevant traits, including Big-5 personality traits (Masland et al., 2015; Riso et al., 1996), maladaptive personality traits such as dependency, mistrust, paranoia, low self-esteem, and schizotypal personality traits (Masland, Drabu, & Hooley, *in press*; Riso et al., 1996), and behavioral tendencies like rumination or social anxiety (Masland et al., *in press).* Findings are mixed with regard to depressive and anxiety symptomatology: Some studies find small to medium correlations (rs =.18 - .38; Chambless & Blake, 2009; Peterson-Post, Rhoades, Stanley, & Markman, 2014; Smith & Peterson, 2008; White, Strong, & Chambless, 1998), and others do not (Hooley et al., 2012; Hooley & Teasdale, 1989; Masland et al., *in press*, 2015). Because PC is a single-item measure, null findings must be evaluated cautiously. Single item measures

typically demonstrate limited reliability and do not capture as much variance as multi-item scales. This reduces statistical power and, thus, the likelihood of finding significant correlations, particularly if sample sizes are small, and also if there is a true, albeit weak, correlation.

Furthermore, there is some heterogeneity in the administration of the PCM across these studies. Some of these studies asked respondents to rate PC from a relative with whom they lived—typically a spouse or parent, whereas others simply allowed respondents to rate PC from their most influential relationship. We know little about how people form PC judgments for different relationships, and it is possible that PC ratings for different targets may have different correlates. Riso and colleagues (1996), for instance, found that whereas spousal PC ratings were correlated with self-report scales referring to their marital relationship, familial PC ratings were related to more general feelings of social belonging. Similarly, Renshaw (2007) found that PC ratings for parents, friends, and romantic partners did not individually predict increases in depressive symptoms among college students. However, when selecting only PC ratings for people with whom participants lived, PC did predict changes in depression. It is possible, then, that in studies that allow participants to provide a PC ratings for any influential person, heterogeneity in participants' choices of relationships may hide subtle but systematic variation in how PC correlates with different types of factors, depending on the target of the PC rating.

Present study

The extant research provides evidence of both relationship-specific and individual-trait components of the PCM. However, the extent to which each of these components is represented in a PC rating is unclear. In the present study, we shed light on this question by evaluating the relationships between PC ratings for 5 different relationships and a variety of relationship-specific and individual variables. Relationship-specific variables are considered those that assess perceptions, descriptions, and evaluations of specific relationships, such as relationship satisfaction with a romantic partner or quality of bonding with parents. Individual variables are considered to be those that assess a person's general cognitive, affective, and

17

behavioral tendencies (e.g., symptomatology, personality, cognitive biases), as well as more general social appraisals (e.g., global perceived social support). Participants included married or dating individuals recruited from Amazon's MTurk platform. Participants completed a PCM for 5 different relationship targets followed by a battery of questionnaires assessing individual- and relationship-level variables.

We conducted four sets of analyses. First, we examined the similarity of PC ratings across respondents' 5 relationships. Second, we estimated the correlations between each PC rating and several clinically-relevant individual traits and relationship characteristics. We also estimated the correlations between each person's average PC rating and these variables. Lastly, we used backwards linear regression to evaluate which correlates best accounted for variance in PC ratings for mothers and partners, the two relationships most commonly selected as being "most important". Though we hypothesized that PC would demonstrate patterns of correlations consistent with its specificity to each relationship, we also hypothesized that PC would be somewhat consistent across a person's many relationships. Furthermore, we hypothesized that this consistency, operationalized as a person's average PC, would be related to individual-level cognitive and affective factors.

Methods

Participants and Design

Participants included people in dating or married relationships recruited from Amazon's Mechanical Turk (MTurk). MTurk is an online platform that allows individuals ("workers") with internet access to complete for-pay tasks posted by "requesters". Workers may browse tasks and choose ones they would like to complete, and requestors can control which workers can see advertisements for their tasks. Numerous studies have established the reliability and validity of data collection on MTurk for both behavioral tasks and self-report surveys that assess a range of psychological constructs (e.g., Behrend, Sharek, Meade, & Wiebe, 2011; Buhrmester, Kwang, & Gosling, 2011; Crump, McDonnell, & Gureckis, 2013).

In the current study, interested MTurk workers completed a screening questionnaire for "a psychology study involving answering questions about your personality, family life, and relationships." Before responding, workers provided informed consent. The screening questionnaire assessed our primary inclusion criteria: participants were required to be in dating or married relationship, over 18 years of age, and speak English as their primary language. The screener was displayed only to workers meeting the following criteria: located in the United States, completed at least 50 prior MTurk jobs ("HITS"), and having received an approval rate of at least 90% for previously completed HITS. These latter two criteria help to ensure that workers completing the survey had a history of providing acceptable work. Workers who completed the screener were paid $0.10. Of the 517 workers who completed the screener 348 met our inclusion criteria and were given access to the study questionnaires.

Of the eligible workers, 197 people chose to participate in the study ("participants" from here forward). These participants were compensated $3.00 for completing the study questionnaires. Sixty percent of participants were women (n = 119). The average age of the sample was 35.2 years (SD = 9.9). Sixty-eight percent of participants were married, and the remaining 32 percent were dating. Ninety percent reported being in opposite-sex relationships, 10 percent in same-sex relationships, and 1 respondent described their relationship as "other". Participants had been in their relationships on average for 9.26 years (SD = 8.4). Participants mostly lived with their partners (85%), though another 10 percent of participants were 'somewhat' cohabitating. The remaining 5 percent were not cohabitating. Examining race, 82 percent of participants were White, 3 percent African American, 6.6 percent Hispanic, 4 percent were Asian, 2 percent were Native American, and 2 percent were of mixed race. With respect to sex, race, education, occupation, or relationship status, participants (n = 197) were not significantly different from those eligible workers who chose not to participate (n = 151). However, there was a trend towards participants being slightly older (M = 35.2 years) than workers (M = 33.2 years), $t(334.27) = 1.93$, $p = .055$.

One concern when conducting studies on MTurk is the possibility that automated bots might complete the questionnaire. Several steps were taken to minimize this risk: The screener was displayed only to participants having a history of approved work; the survey was listed under a different title, which was given only to eligible workers after completing the screener; and eligible workers were given a unique eligibility code after completing the screener, which needed to be copied from the screener and entered into the survey in order for the responses to be approved; the screener and survey included numerous Captcha responses to ensure respondents were human; and lastly, any survey that was completed in fewer than 25 minutes was excluded.

Individual Measures

Perceived Criticism Measure (PCM; Hooley & Teasdale, 1989). The PCM was used to assess PC from 5 targets: romantic partners, mothers (or mother-figures), fathers (or father-figures), best friends, and a 5th person of each participant's choosing. For each target, participants responded to the question "How critical is your [*relative/partner, etc.*] of you?" on a 1 – 10 scale (not at all critical - very critical).

Dysfunctional Attitudes Scale – Form A (DAS-A; Weissman & Beck, 1978). The DAS-A assesses the depressogenic cognitive schema thought to underlie vulnerability to depression according to Beck's cognitive model of depression (Beck, 1967). High DAS-A scores predict increases in depression symptoms after stressful life events (e.g., Joiner, Metalsky, Lew, & Klocek, 1999). Participants rate their agreement with 40 statements (e.g., "If I fail partly, it is as bad as being a complete failure") on a 7-point Likert scale. All items were summed to provide an index of dysfunctional attitudes. Higher scores indicate stronger dysfunctional attitudes. The DAS-A demonstrates satisfactory internal consistency and test-retest reliability in a variety of populations (Dobson & Breiter, 1983; Oliver & Baumgart, 1985; Weissman, 1980). Cronbach's alpha for all scales in this study can be found in Table 1.

Social Provisions Scale (SPS; Cutrona & Russell, 1987). The SPS assesses perceived availability of social support (commonly referred to as "perceived support"). Perceived support has been conceptualized as cognitive personality trait (Lakey & Cassady, 1990) and is consistently associated with greater psychological well-being and fewer mental health problems (Cohen & Wills, 1985; Ibarra-Rovillard & Kuiper, 2011). Participants rated their agreement with 24 statements on a 1 – 4 scale (strongly disagree – strongly agree; e.g., "There are people I can depend on to help me if I really need it."). All items were summed to provide a global measure of perceived support. Higher scores indicate greater perceived support. The global SPS score demonstrates excellent internal reliability (α =.92; Cutrona & Russell, 1987).

Experiences in Close Relationships – Relationship Structures (ECR-RS; Fraley et al., 2015). The ECR-RS is a brief measure of attachment anxiety and avoidance. It can be used to assess attachment insecurity in specific relationships (e.g., Fraley, Heffernan, Vicary, & Brumbaugh, 2011) or globally (e.g., Fraley et al., 2015). In this study, we used the ECR-RS to assess *global* attachment insecurity. Participants rated their agreement with 9 statements (e.g., "It helps to turn to a person in times of need.") on a 1-7 scale (strongly disagree – strongly agree). For simplicity, we combined the anxiety and avoidance subscales to provide a general index of attachment insecurity. Higher scores indicate greater attachment insecurity. Correlations between global scores and relationship-specific scores range from r = .27 - .61, and internal consistency of the global scale is equal to or greater than α = .81 (Fraley et al., 2015).

Brief Symptom Inventory – 18 (BSI-18; Derogatis, 2000). The BSI assesses 3 domains of clinical symptoms: somatization, depression, and anxiety. On a 0-4 scale (not at all – extremely), participants rated the extent to which they had been distressed by 18 symptoms in the last 7 days. An average raw score was computed for all 18 symptoms to provide a global symptomatology score. Higher scores indicate greater global symptomatology. The BSI-18 has

been shown to have high internal consistency and modest 15-day test-retest reliability (Andreu et al., 2008; Derogatis, 2000)

Scale for Positive and Negative Experiences (SPANE; Diener et al., 2010). The SPANE assesses positive and negative emotionality in the previous month. Participants rated how much they experienced 6 positive (good, pleasant, happy, positive, joyful, contented) and 6 negative (bad, unpleasant, negative, sad, afraid, angry) feelings over the past 30 days. These feelings were rated on a 5-point Likert scale from 1 to 5 (very rarely or never - very often or always). Items from each subscale were summed. Higher scores indicate greater positive and negative feelings. The SPANE demonstrates adequate internal consistency α = .81 - .89 and 1-month test-retest reliability r = .63 - .68.

Parental-relationship Measures

Childhood Trauma Questionnaire (CTQ; Bernstein, Ahluvalia, Pogge, & Handelsman, 1997). The CTQ assesses 5 types of maltreatment experienced in the home environment: emotional abuse and neglect, physical abuse and neglect, and sexual abuse. Maltreatment is known to create long lasting changes in neural development and functioning (Heim & Binder, 2012), increasing risk for poor mental health (Edwards, Holden, Felitti, & Anda, 2003). Participants rated their agreement with 25 statements (e.g., "My parents were too drunk or high to take care of the family") on a 5-point scale from "never true" to "very often true". We conceptualized the CTQ as reflecting poor quality parental relationships. To reduce the number of scales presented in our correlation tables, physical and emotional abuse were combined into a single 'abuse' variable (α = .93), and physical and emotional neglect were combined into a single 'neglect' variable (α = .93).

Parental Bonding Instrument (PBI; Parker, Tupling, & Brown, 1979). The PBI is an assessment of parental relationships. It assesses 2 dimensions of childhood bonding with parents: care/warmth and overprotectiveness/control. Participants rated the extent to which 25 statements described their mothers and fathers (e.g., My mother [or father] "was affectionate to

22

me."). Response choices included "very like", "moderately like", "moderately unlike", and "very unlike". Mothers and fathers were rated separately. The PBI is reliable over 20 years (rs = .59 - .75), suggesting that mood, pathology, and life experiences may not influence perceptions of parental bonding once they have formed (Wilhelm, Niven, Parker, & Hadzi-Pavlovic, 2005).

Romantic-relationship Measures

Experiences in Close Relationships – Revised (ECR-R; Fraley, Waller, & Brennan, 2000). The ECR-R assesses anxious and avoidant attachment representations in romantic relationships. Participants rated their agreement with 36 statements (e.g., "I'm afraid that I will lose my partner's love.") on a 1– 7 scale (strongly disagree - strongly agree). Participants were instructed to respond to each item by thinking about how they generally feel in their "emotionally intimate relationships" – not just in their current romantic relationship. Higher scores represent more insecure attachment. The ECR-R demonstrates high temporal stability and good specificity, as it explains 30-40% of the variance in attachment-related emotions towards a romantic partner, but only 5-15% of the variance in attachment-related emotions towards family and friends (Sibley, Fischer, & Liu, 2005).

Attributions of Criticism Scale (ACS; Allred & Chambless, 2014). The ACS is a 22-item self-report inventory assessing individuals' explanations for romantic partners' criticism. Respondents were given the prompt "When your relationship partner criticizes you, to what extent do you believe he/she…" They then rated 9 negative attributions about inflicting harm (e.g., "…is trying to hurt you") and 13 positive attributions about fostering growth (e.g., "…has your best your best interest at heart") on a 5-point scale (not at all – completely). The two factors are slightly correlated (r = -.17), and both have been correlated with PC (Allred & Chambless, 2014). Alpha reliabilities for both subscales range between .91 and .93 in Caucasian and African American samples (Allred & Chambless, 2014).

Couples Satisfaction Index (CSI; Funk & Rogge, 2007). The CSI assesses relationship satisfaction in respondents' current romantic relationship. Participants provided

ratings for 16 questions on a 0 – 5 scale, where higher numbers indicated greater satisfaction. The CSI demonstrates strong correlations with other widely used measures of relationship satisfaction, such as the Dyadic Adjustment Scale (r = .89) but contains less measurement noise because it focuses on subjective evaluations (e.g., "I have a warm and comfortable relationship with my partner.") and excludes questions about interaction patterns (e.g., "How often do you or your mate leave the house after a fight?").

Positive-Negative Relationship Quality-4 (PNRQ-4; Fincham & Rogge, 2010). The PNRQ-4 assesses positive and negative dimensions of relationship quality in romantic relationships, which are theorized to be separate dimensions. For the negative qualities, participants are instructed "Considering only the negative qualities of your [romantic] relationship and IGNORING the positive ones, evaluate your relationship on the following qualities: bad, miserable, empty, and lifeless." Participants rated each of the 4 adjectives on a 7-point Likert scale from "not at all" to "extremely". Participants rated 4 positive relationship descriptors in a comparable manner: enjoyable, pleasant, strong, and alive. Studies indicate that the PNRQ-4 demonstrates incremental validity with respect to the CSI (Fincham & Rogge, 2010).

Results

Descriptive statistics are presented in Table 1. Participants reported the highest PC from mothers and the lowest from friends. In their PC ratings, participants most commonly identified their partner as their most important relationship (75%), followed by mother (11%), child (4.6%), friend (4%), father (2.5%), and sibling (2.5%). The 'other' PC category mostly consisted of relatives: 36 percent rated siblings; 29 percent rated a child; 16 percent rated a cousin, aunt, uncle, godparent, or grandparent; 19 percent rated a friend or someone else. Mean scores for the DAS-A were slightly higher than those for college students (M = 113, SD not provided; Weissman & Beck, 1978), but lower than those for a large sample from the general population (M = 137.8, SD = 23.6; de Graaf, Roelofs, & Huibers, 2009). The mean raw BSI – Global score

Table 1. Descriptive statistics of self-report scales

	n	Mean	SD	Median	Min	Max	α
Perceived Criticism							
Average	197	3.39	1.72	3	1	9.2	0.72
Mother	176	4.18	2.76	3	1	10	
Father	163	3.45	2.69	2	1	10	
Partner	197	3.61	2.60	3	1	10	
Friend	192	2.78	2.02	2	1	9	
Other	195	3.12	2.44	2	1	10	
Individual Factors							
DAS-A	197	119.57	39.62	117	40	251	0.95
ECR-RS	196	25.13	12.45	23	9	63	0.92
SPS – Total	180	79.16	11.70	80	38	96	0.95
BSI – Global	197	0.49	0.62	0.28	0	2.94	0.95
SPANE – Negative	197	11.46	5.09	10	6	30	0.93
SPANE – Positive	196	23.70	5.27	24	6	30	0.94
Family/Parent Factors							
CTQ – Emot. Abuse	193	9.19	5.28	7	5	25	0.93
CTQ – Emot. Neglect	193	9.99	5.27	8	5	25	0.95
CTQ – Phys. Abuse	194	7.29	3.97	5	5	23	0.87
CTQ – Phys. Neglect	192	7.30	3.44	6	5	20	0.83
CTQ – Sexual Abuse	192	6.59	3.93	5	5	25	0.94
PBI – Father – Care	177	21.79	10.20	24	0	36	0.96
PBI – Father – OP	178	11.98	8.30	11	0	38	0.91
PBI – Mother – Care	190	25.16	10.26	29	0	36	0.96
PBI – Mother – OP	188	14.84	9.08	14	0	39	0.91
Partner Factors							
ACS – Negative	194	17.63	8.62	14	9	45	0.94
ACS – Positive	192	43.32	11.67	44.5	13	65	0.93
CSI	189	66.50	14.88	71	0	81	0.98
ECR-R – Anxiety	194	2.38	1.36	2.00	1	6.8	0.97
ECR-R – Avoidance	195	2.85	1.16	2.56	1	6	0.97
PNRQ – Negative	197	8.59	5.30	7	4	28	0.95
PNRQ – Positive	197	24.01	4.67	25	4	28	0.96

Note. DAS-A – Dysfunctional Attitudes Scale – Form A; ECR-RS – Experiences In Close Relationships – Relationship Structures; SPS – Social Provisions Scale; BSI – Brief Symptom Inventory; SPANE – Scale for Positive and Negative Experiences; CTQ – Childhood Trauma Questionnaire; PBI – Parental Bonding Instrument; ACS – Attributions of Criticism Scale; CSI – Couples Satisfaction Index; ECR-R – Experiences in Close Relationships – Revised; PNRQ – Positive/Negative Relationship Quality

was somewhat higher than that reported for non-patients (M = .30, SD = .31; Derogatis, 2000) but substantially lower than that reported for psychiatric outpatients (M = 1.32, SD = .72). Compared to a large college sample (Diener et al., 2010), participants reported less frequent negative moods but similar positive moods in the past 30 days (from Diener et al. (2010): Mean negative mood = 15.36 (3.95); Mean positive mood = 22.05 (3.73)).

Correlations among PC ratings

First, we investigated the extent to which people perceive similar amounts of criticism across 5 close relationships (see Table 2). All PC ratings were significantly correlated, indicating that PC ratings are similar across a person's relationships. However, there was some variation in the magnitudes of these relationships. The strongest relationships were observed between PC ratings for friends, partners, and 'other'. This is consistent with previous findings suggesting that non-parental relationships have a high similarity in PC (White et al., 1998). The weakest relationship was observed between PC ratings for fathers and 'other'. Although the correlations among PC ratings fall into the small-medium range, that these ratings are robustly linked across relationships suggests the possibility of a shared underlying factor or even an individual-specific 'baseline' level of PC. In order to explore this idea in subsequent correlation and regression analyses, we computed an average PC score for each person. Internal reliability was acceptable (Cronbach's α = .72).

Table 2. Correlations between perceived criticism ratings for 5 relationships

	Partner	Mother	Father	Friend
Partner	-			
Mother	0.37***			
Father	0.25**	0.38***		
Friend	0.42***	0.34***	0.36***	
Other	0.41***	0.35***	0.21**	0.41***

†*p*<.10, **p*<.05, ***p*<.01, ****p*<.001

Correlates of PC Ratings by Relationship

Next, we examined the correlations between each PC rating, including average PC, and several individual and relationship variables. Table 3 *Table 3*displays the significant correlations ($p < .05$) between PC ratings and the individual, parental, and partner variables. Significance was determined after a Holm-Bonferroni correction for multiple comparisons, which is less conservative than a traditional Bonferroni correction but adequately reduces Type I error (Holm, 1979).

With the exception of positive moods, all individual variables were related to PC ratings for mothers, fathers, and partners. Only dysfunctional attitudes and global symptoms were correlated with PC ratings for friends and 'other'. Positive mood demonstrated the weakest and most inconsistent relationships with PC ratings, as it was correlated with only partner PC ratings. Individuals' average PC demonstrated consistently larger correlations with all individual factors. This may be a function of the increased reliability of average PC. Average PC was related most strongly to global symptoms and dysfunctional attitudes.

Among parental-relationship constructs, factors demonstrated substantial specificity in their correlations with mother- and father-PC ratings. PC ratings for mothers demonstrated medium-to-strong correlations with Mother-Care and -Overprotection scales and small-to-medium correlations with abuse, neglect, and Father-Overprotection. PC ratings for fathers correlated most strongly with Father-Care (positively) and correlated moderately with abuse and Father-Overprotection (both negatively). Though average PC ratings correlated with many of these family variables, these correlations may be driven mainly by parental PC, as average PC ratings for non-parents were not correlated with any of these scales. These findings suggest that PC ratings for non-parental relationships may not be associated with parental relationship quality or traumatic history in one's childhood home.

Table 3. Correlations between PC ratings and Individual, Parental, and Partner Factors

	PC Ratings						
Individual Factors	Mother	Father	Partner	Friend	Other	Average	
Dysfunctional Attitudes	**0.32***	**0.35***	**0.24***	**0.26****	**0.22***	**0.39***	
Attachment Security	**0.24***	**0.26***	**0.27****	0.14	0.14	**0.30***	
Perceived Support	**-0.25***	**-0.25***	**-0.28****	-0.15	-0.19	**-0.32***	
BSI - Global	**0.28****	**0.34***	**0.28****	**0.24***	**0.21***	**0.40***	
SPANE - Negative	**0.26***	**0.28****	**0.34***	0.14	0.2	**0.36***	
SPANE - Positive	-0.18	-0.18	**-0.22***	-0.05	-0.07	**-0.22***	
Parental Factors							Average Excluding Parents
CTQ - Neglect	**0.26***	0.23	0.08	0.04	0.07	0.21	0.09
CTQ - Physical Abuse	**0.33***	**0.29***	0.17	0.20	0.17	**0.32***	0.23
CTQ - Sexual Abuse	0.22	0.01	0.13	0.07	0.17	0.17	0.17
PBI Mother - Care	**-0.43***	-0.12	-0.18	-0.10	-0.11	**-0.27****	-0.17
PBI Mother - Overprotection	**0.52***	0.03	0.14	0.11	0.07	**0.24***	0.13
PBI Father - Care	-0.18	**-0.47***	-0.10	-0.07	-0.14	**-0.27***	-0.13
PBI Father - Overprotection	**0.28***	**0.28***	0.08	0.17	0.10	**0.24***	0.13
Partner Factors							Average Excluding Partners
CSI	-0.18	-0.12	**-0.36***	-0.06	-0.18	**-0.27****	-0.19
PNRQ - Positive	-0.08	-0.02	**-0.31***	0.00	-0.16	-0.18	-0.10
PNRQ - Negative	**0.24***	0.14	**0.44***	**0.24***	**0.22***	**0.37***	**0.28****
ECR-R Avoidance	**0.28****	**0.29****	**0.29****	0.19	0.19	**0.37***	**0.33***
ECR-R Anxiety	**0.24***	0.24	**0.28****	0.06	0.18	0.28**	**0.23***
ACS - Positive	-0.07	0.00	-0.17	0.03	-0.04	-0.09	-0.04
ACS - Negative	**0.27***	0.17	**0.43***	**0.24***	0.28**	**0.41***	**0.33***

[†]$p<.10$, *$p<.05$, **$p<.01$, ***$p<.001$

Note: Significant correlations are shown in bold.

As might be expected, partner-relationship constructs were most strongly and consistently correlated with PC ratings for partners. Relationship satisfaction (CSI), for instance, loaded exclusively on partner PC. However, in comparison to parental variables, romantic-relationship variables demonstrated much less specificity in their relationships with PC ratings. For instance, PNRQ-Negative and ACS-Negative were correlated with PC ratings for all targets except fathers. Romantic attachment (both avoidance and anxiety) were also related to mother and father PC ratings. And in further contrast to the family/parent factors, average PC ratings were correlated with several romantic-relationship factors even when the average excluded PC ratings for partners. This pattern of results suggests that PC ratings may, to some extent, reflect more global or non-specific relationship patterns and experiences.

Which variables best predict PC?

Many of the individual and relationship variables included in this study represent overlapping or related constructs. The purpose of our final set of analyses was to identify which factors account for the majority of variance in partner, mother, and average PC. For each PC rating, we selected the variables with which it was most strongly correlated. Then we submitted these variables to a stepwise backward regression using the stepAIC function from the *MASS* package (Venables & Ripley, 2002) in R (R Core Team, 2014). Variable removal at each step was determined by the aikake information criterion (AIC), which balances parsimony and goodness-of-fit for each potential model.

For our initial model of partner-PC, we selected the following 6 variables: CSI, PNRQ – negative, ACS – negative, BSI – global, SPANE – negative, and SPS. Given the possibility that PC could itself be an individual difference, we also included an average PC variable that excluded partner ratings. The final model included 3 variables and accounted for 36% of the variance in partner PC ratings, $F_{3, 190} = 35.80$, $p < .001$. Table 3 displays the standardized regression coefficients for the final model. Average PC had the largest effect, followed by negative relationship experiences and negative attributions of partner criticism.

29

For our initial model of mother-PC, we selected the following 8 variables: CTQ - physical abuse, CTQ - emotional abuse, PBI - maternal overprotection and care, BAS – global, DAS, ECR-R avoidance, and average PC (excluding mother-PC). The final model included 3 variables (see Table 4) and accounted for 46% of the variance in mother-PC ratings $F_{4, 163} =$ 46.16, $p < .001$. Maternal overprotection demonstrated the largest effect, followed by average PC and maternal care.

For our initial model of average-PC, we selected the following 9 variables: PNRQ – Negative, ACS – Negative, BSI – Global, DAS, ECR-R Avoidance, SPS, CTQ – Physical Abuse, PBI Mother Care and Overprotection. The final model included 4 variables (see Table 3) and accounted for 29% of the variance in average PC ratings, $F_{4, 186} = 19.11$, $p < .001$. Dysfunctional attitudes exerted the largest effect, followed by physical abuse, negative attributions of partner criticism, and negative experiences in a current romantic relationship.

Table 4. Final models predicting Partner PC and Mother PC

	ß	SD	t	p
Partner PC				
PC Average	0.33	0.06	5.37	0.000
PNRQ - Negative	0.27	0.07	3.92	0.000
ACS - Negative	0.19	0.07	2.81	0.006
Mother PC				
PBI - Mother - OP	0.43	0.07	6.08	0.000
PC Average	0.39	0.06	6.78	0.000
PBI - Mother - Care	-0.15	0.07	-2.24	0.027
Average PC				
DAS-A	0.22	0.07	3.09	0.002
CTQ - Physical Abuse	0.20	0.07	3.00	0.003
ACS - Negative	0.20	0.08	2.60	0.010
PNRQ - Negative	0.17	0.08	2.06	0.041

Discussion

The purpose of the present study was to examine the extent to which the PC construct assesses individual versus relationship processes. Consistent with previous studies, we found clear evidence that PC is in part relationship-specific (Riso et al., 1996; White et al., 1998). Relationship variables were correlated most strongly, and often exclusively, with PC ratings for the same target relationship. For instance, the mother subscales of the PBI correlated most highly with mother-PC; and relationship satisfaction with one's romantic partner correlated most highly with partner-PC. Riso and colleagues (1996) similarly found that scales assessing marital quality were related to spousal PC but not family PC. Furthermore, in our regression analyses, relationship-specific variables accounted for more variance in mother-PC and partner-PC than individual difference variables like dysfunctional attitudes, symptomatology, or negative affect. These findings indicate that PC ratings do indeed assess negative experiences and attitudes specific to each relationship. This finding also lends support to the prevailing hypothesis that PC predicts clinical outcomes because it assesses harmful relationship experiences, such as criticism and hostility.

However, our analyses also suggest that PC provides a window into several individual-level processes. These may include cognitive vulnerabilities such as dysfunctional attitudes, attachment insecurity, and low perceived support; and affective vulnerabilities such as negative affect or global symptomatology. The presence of these vulnerabilities among people who perceive high levels of criticism may be a key reason why they are more likely to experience poor clinical outcomes. Our findings are consistent with a number of previous studies. For instance, Masland and colleagues (*in* press) also found that PC was linked to attachment insecurity, as assessed by the Revised Adult Attachment Scale, in a large community sample of MTurk participants. Furthermore, Riso and colleagues (1996) demonstrated that PC from one's family (generally, and not in reference to a particular relative) was related to perceived belonging support—the belief that one has friends or family with whom they can enjoy

31

meaningful activities. However, they did not find perceived belonging support to be related to spousal PC. It is worth noting that Riso and colleagues (1996) and Masland and colleagues (*in press*) both failed to find any relationships between PC and a wide array of personality variables such as neuroticism, extroversion, and personality disorder traits (e.g., paranoid, dependent, avoidant). Thus, although PC may be influenced by some individual-level factors, it does not appear to be tied with personality. However, these studies also used single-item measures of PC. In the present study, a person's Average PC across 5 relationships was much more robustly correlated with individual factors than any single PC rating. This may be because combining multiple PC ratings increases the variance in PC across the sample, as well as the reliability of the measure. It may also wash out the relationship-specific influences on PC, thus indexing how much criticism a person *generally* tends to perceive in his or her important relationships. It is also worth noting that the correlations found in this study, though often statistically significant, were still modest in magnitude. This is consistent with a comprehensive review of the PC construct by Renshaw (2008). Thus, though PC may be linked with some individual level processes or constructs, it is not redundant with them. Furthermore, these relationships may only be statistically detectible in larger samples.

With regard to symptomatology, our findings concur with those prior studies that have also found associations with aspects of psychopathology, such as global symptomatology (Chambless & Blake, 2009), global functioning (Chambless et al., 2001), and depressive symptomatology (White et al., 1998). However, our findings are at odds with some studies that have failed to find a relationship with psychopathology, such as Hooley and Teasdale (1989), Masland and colleagues (2015), and Masland and colleagues (*in press*), all of whom measured depressive symptomatology using the BDI. There is not yet a clear pattern to explain these conflicting results. It has been hypothesized that perhaps correlations between PC and symptomatology are more likely to be found in community or non-clinical samples. This does not appear to be the case, as the two groups of studies cited above are each composed of two

studies with non-clinical samples and one study with a clinical sample. An alternative possibility is that PC may be more reliably correlated with more global measures of functioning or symptomatology, which assess social experiences in addition to personal thoughts, feelings, and behaviors. PC is an inherently social construct, so it would make sense for it to be more consistently related to measures that assess a person's social experiences. A final consideration is sample size. It may be that PC has only a modest relationship with symptomatology and that this can be consistently captured in only studies with larger sample sizes. Riso and colleagues found a correlation of .24 between spousal PC and the Hamilton Rating Scale for Depression. In the present study, Friend-PC and Other-PC correlated with global symptomatology at similar magnitudes (.24 and .21 respectively), which was statistically significant likely due to our large sample size. Similarly, Masland and colleagues (2015) compared BDI across high- (n = 41) and low-PC (n = 35) groups based on a median split, and found a non-significant difference (p=.10), but a small-to-medium effect size (d = .38) indicating that a larger sample may have boosted power enough to make the difference statistically significant.

In addition to individual factors, PC may also reveal an individual's relationship patterns. Our analyses indicate that people perceive similar amounts of criticism in their relationships. We also observed that criticism perceived from parents, friends, and 'others' was related to several aspects of a person's romantic relationship. These findings suggest that individuals may have similar experiences in their relationships and that PC ratings can reveal some of these patterns. Thus, high PC from one's romantic partner may predict clinical outcomes not only via negative experiences in *that* relationship, but also via criticism and negative experiences in several other relationships. PC, in this way, may provide insight into the quality of person's overall social environment.

The substantial correlations among PC ratings also raise the possibility that PC is to some extent an individual difference variable. In both our correlation and regression analyses,

PC ratings from non-target relationships (e.g., mother-PC) accounted for as much variance in a target PC rating (e.g., partner PC) as relationship-specific variables. In fact, average PC was the strongest predictor of partner-PC, followed then by relationship-specific constructs. These analyses illustrate PC's simultaneous specificity and generality. Accurate models of the PC construct must account for this duality. One possibility is that people have a trait-like 'baseline' for PC that is moderated by specific relationship experiences and characteristics. Our regression analysis on average PC suggests a few processes that may account for a person's 'baseline' PC, including dysfunctional attitudes, negative attributions of criticism, and more distally, childhood abuse. Childhood abuse may indirectly shape PC later in life, as it is associated with more problematic dysfunctional attitudes and cognitive biases (Wells, Vanderlind, Selby, & Beevers, 2014). 'Baseline' PC may also be driven by genuine similarities in an individual's relationships: People may select into more critical relationships or behave in ways that elicit more criticism from others (Gerlsma et al., 2009; Peterson et al., 2009). This possibility is supported by our finding that negative relationship quality in romantic relationships predicted average PC.

There are important limitations to this research. Because these data are cross-sectional, we cannot make inferences about how or why PC is related to the various constructs examined here. However, the goal of this study was not to characterize the temporal or causal relationships between these variables but to understand the overall structure of PC and define the range of information that it can provide. Understanding the structure and correlates of PC can shape future studies' hypotheses about the origins and consequences of PC. Another limitation is the potential for inflated correlations among PC ratings due to shared method variance and the proximity with which participants responded to these questions. However, the correlations found here are consistent with those found by White and colleagues (1998), who assessed undergraduates' PC ratings for parents and peers. Future studies may rule this problem out more precisely by using multiple measures of perceived criticism or by having

participants provide PC ratings at different time points. Another key limitation of this study is the composition of the sample. Because we employed a community sample from MTurk, these findings may not generalize to PC assessed among inpatients or outpatients, for whom PC is most important. However, numerous studies on PC have also used community and non-clinical samples (e.g., Chambless & Blake, 2009; Masland et al., *in press*, 2015; Peterson & Smith, 2010) and findings from these studies are not inconsistent with studies of clinical samples. Relatedly, because this sample included relatively few ethnic minorities, these findings may also not generalize to non-white populations. Prior research on PC suggests that the meaning of the PC construct may differ across cultures (Allred & Chambless, 2014). In order to address these issues, future research should examine differences in PC's correlational network across different ethnic and clinical groups.

PC is a rich and complex construct. Our analyses indicate that we must expand our conceptualization of PC beyond criticism in a target relationship. PC may provide as much information about a relationship as it does about the person providing the rating: People reporting high PC in one relationship also report higher PC in other relationships; more negative experiences in romantic relationships; as well as a constellation of clinical risk factors like dysfunctional attitudes, insecure attachment, and low perceived social support. This myriad information present in a PC rating may account for its ability to predict clinical outcomes reliably and incrementally over more circumscribed measures. Its simultaneous simplicity and richness make PC a valuable clinical tool.

Paper 2: Global perceived criticism ratings incrementally predict daily social appraisals over daily criticism exposure

Steven A. M. Felix, Jill M. Hooley, & Christine I. Hooker

Manuscript under revision

Abstract

      It is generally assumed that a cross-sectional assessment of perceived criticism ("global" PC) predicts clinical outcomes because it approximates person's more granular, day-to-day experiences of criticism from a relative. In the present study we test this assumption by examining whether *daily* assessments of PC account for the relationship between *global* PC and daily well-being. Twenty nine couples from the community completed measures of global PC, relationship satisfaction, and depression. They then responded individually to an online diary questionnaire every evening for 21 days. The diary assessed daily PC, mood, well-being appraisals, and social appraisals. Regression models were conducted in a step-wise manner in order to assess global PC's incremental prediction of daily response variables over and above daily PC, depression, and relationship satisfaction. Higher global PC predicted reduced daily well-being, perceived support, social connection, and perceived partner understanding. However, after controlling for symptomatology, relationship satisfaction, and daily criticism, global PC remained a significant predictor of only social-appraisal variables. These results indicate that PC uniquely reveals how individuals experience and evaluate their social environment on a daily basis. Global PC may predict clinical outcomes not because it approximates daily criticism exposure but because it reveals a bias towards feeling less accepted, understood, and connected to others more generally.

Introduction

Perceived criticism (PC) refers to the amount of criticism perceived globally in one's closest relationship. The Perceived Criticism Measure (PCM; Hooley & Teasdale, 1989) is the most common measure of PC. It is a single-item assessment: Respondents first identify the person with whom they have the closest relationship; they then rate how critical that person is of them on a 10-point Likert scale (Not at all - Extremely). PC is distinct from objective measures of criticism such as that derived from the Camberwell Family Interview (CFI; Leff & Vaughn, 1985) or the Kategoriensystem für Partnerschaftliche Interaktion ([Interaction Rating System] KPI; Hahlweg & Conrad, 1983).

PC is primarily a clinical construct intended to identify psychiatric patients who are at greater risk of poor clinical outcomes (Masland & Hooley, 2015). It robustly predicts symptom fluctuations, relapse, and treatment outcomes across a range of psychopathologies (Chambless & Steketee, 1999; Fals-Stewart et al., 2001; Hooley & Teasdale, 1989; Kwon et al., 2006; Lee et al., 2014; Scott et al., 2012). PC may also impact sub-clinical aspects of well-being, as it predicts increases in depressive symptoms in community samples (Renshaw, 2007; White et al., 1998). A key question facing researchers of PC is why such a simple measure predicts a wide assortment of outcomes. In the present study, we examine the most widely held hypothesis: That PC predicts clinical outcomes because it assesses exposure to criticism and, more specifically, indicates "how much criticism is getting through" (Hooley & Teasdale, 1989, p. 234). We also discuss an alternative hypothesis: That PC acts as a more general indicator of how a person experiences and evaluates many social relationships.

PC and Criticism Exposure

PC is typically conceptualized as a measure of criticism exposure, and many studies have investigated its validity in this respect. PC is reliably correlated with third-party ratings of criticism, such as spouses' own perceptions of how much criticism they express (Chambless & Blake, 2009; Chambless et al., 2001; Smith & Peterson, 2008) and naive raters' overall

impressions of criticism in couples' video-recorded interactions (Chambless & Blake, 2009).

Moreover, PC sometimes, though not always, correlates with criticism indices derived from more objective assessments, such as the CFI or the KPI (Gerlsma et al., 2009; Lee et al., 2014). Together, these findings suggest that the perceived criticism measure picks up on real variation in relationship criticism. This validity in turn provides grounds to hypothesize that PC predicts changes in well-being because it indicates the extent to which a person is exposed to criticism. Criticism, particularly destructive criticism, may increase arousal and negative mood (Hooley, 2007), reduce self-resources such as self-efficacy (Baron, 1988), and reinforce self-criticism (Wearden, Tarrier, Barrowclough, Zastowny, & Rahill, 2000). For vulnerable individuals, such as patients suffering from psychological disorders, experiences of criticism may be particularly stressful, reducing daily well-being and, over time, contributing to poor clinical outcomes (Hooley et al., 2009).

Still, there is relatively little evidence that explicitly supports criticism exposure as the mechanism for PC's prediction of clinical outcomes. The strongest evidence comes from a study that prospectively examined whether PC ratings predicted changes in depression symptoms among college students (Renshaw, 2007). The author found that PC predicted increases in depressive symptoms only when the PC rating referred to someone with whom the student lived. This finding suggests that it is not enough merely to have a highly critical friend or relative; one must be *exposed* to the critical person, or to a generally hostile and stressful home environment, frequently or daily in order to impact well-being.

Somewhat indirect evidence for the "exposure hypothesis" is also provided by Hooley, Siegle, and Gruber (2012) via *f*MRI. These researchers exposed depressed, recovered depressed, and healthy controls to personally-relevant criticism from their mothers during an *f*MRI brain scan. Because PC was found to be unrelated to depression status, they separated the sample into high and low PC groups based on a median split. Their analyses indicated that the high PC group demonstrated reduced activity in the dorsolateral prefrontal cortex (DLPFC)

and greater activity in the amygdala while listening to criticism from their mothers. This is interesting because these neural regions have been associated with affective reactivity as well as cognitive control of emotion during exposure to emotional stimuli (Hooker et al., 2010; Morrison & Salzman, 2010; Ochsner & Gross, 2005). Furthermore, this pattern of activity has also been demonstrated in remitted-depressed patients (Hooley et al., 2009), suggesting that individuals with high PC may be more likely to experience affective dysregulation in the context of criticism exposure. Thus, this study provides some evidence for the role of criticism exposure in shaping clinical outcomes. However, it does not necessarily imply that PC results from criticism exposure or acts as a proxy for criticism exposure.

PC as a Trans-Relational Lens

Though the exposure hypothesis is intuitively appealing and reasonable, particularly given evidence that the perceived criticism measure taps into genuine criticism, there are important alternatives worthy of consideration. One possibility is that, in addition to describing criticism in a particular relationship, PC may also reveal a person's habitual ways of experiencing and evaluating all of their social relationships. At least two studies have found that people perceive similar levels of criticism in their close relationships (Felix, Masland, & Hooley, 2017; White et al., 1998). In a large community sample, the latter researchers asked participants to provide PC ratings for 5 important relationships: romantic partners, mothers, fathers, best friends, and a self-selected 5th person. All pairs of PC ratings were significantly correlated. The authors also tested the relative strength of these correlations compared to variables that were specific to a target relationship. For instance, in one regression, they modeled PC from a romantic partner as a function of both relationship satisfaction (with the romantic partner) and average PC in the remaining 4 relationships. They found that a person's average PC was a stronger predictor of PC ratings about a romantic partner than was relationship satisfaction. In other words, though a given PC rating may partially reflect relationship-specific experiences or attitudes, such as exposure to criticism *in that relationship*, it

also indicates whether a person has a *general tendency* to perceive criticism in other close relationships. Thus, PC may be a *trans-relational* construct.

Similarity in PC across relationships suggests that PC taps into a trait-like factor that shapes how one experiences and evaluates relationships. Though it is unclear exactly what this factor might be, PC ratings may predict clinical outcomes because they describe whether a person tends to have positive and supportive relationships, or more negative and hostile relationships. Indeed, Felix, Masland, and Hooley (2017) also found that PC ratings were related to global perceived social support, further suggesting that PC is indicative of the quality of a person's overall social environment—not merely a single relationship, and not merely criticism.

Present study

In the present study, we investigated these two competing conceptualizations of PC and their respective explanations for PC's prediction of well-being. Specifically, we examined whether PC's prediction of daily well-being is completely explained by exposure to criticism or, alternatively, whether PC demonstrates incremental predictive validity over criticism exposure.

Twenty-nine couples were recruited from the community to participate in a daily diary study. Both members of each couple were considered separate participants and completed all parts of the study independently of their partners. Couples first completed measures of PC, depression, and relationship satisfaction. To obtain a global measure of criticism exposure, all participants provided a rating for how critical they thought they were of their partners. Subsequently, for 21 consecutive days each participant responded to a daily questionnaire online. To measure perceived daily criticism exposure, participants reported the extent to which they perceived their partners as critical that day ("daily PC"). Furthermore, participants provided daily ratings for several aspects of well-being.

If PC predicts clinical outcomes merely because it indicates the extent of criticism exposure, then PC should demonstrate little, if any, incremental predictive validity when

measures of criticism exposure are statistically controlled. However, if PC operates through processes other than individual experiences of received criticism—for instance, by shaping the ways people perceive events in their relationships—we would expect PC to make an incremental prediction to daily well-being even when daily criticism is controlled.

Material and Methods

Participants

Thirty-eight dating or married couples were recruited from Craigslist. Criteria for inclusion consisted of being over 18 years old and being in a dating relationship for more than 3 months. Because current psychopathology could be a potential confounding variable, all participants were screened using the Mini International Neuropsychiatric Interview (M.I.N.I.), English Version 5.0.0 (Sheehan et al., 1998). Participants were excluded if they demonstrated current psychopathology, as indicated by the M.I.N.I.) or if they anticipated being apart for any longer than a few days over the course of their participation in the study. Four couples were immediately excluded due to the presence of a current psychological disorder. The remaining 34 couples participated in the study. However, four of these couples withdrew before study completion, and one was administratively terminated by the researchers due to poor diary compliance. Among the four couples that withdrew voluntarily, one had concerns about privacy, and the remaining three did not think they could keep up with answering the questionnaires each day. For these analyses, the final sample consisted of 29 couples (27 males, 31 females). Mean age was 24.5 years (SD = 4.05). Mean relationship length was 9.2 years (SD = 7.09), and all couples had been together longer than 1 year. Twenty-six couples were dating, and three were married or engaged. About one-third of couples (n=10) were living together.

Study Procedure

All couples were invited to an initial laboratory session. After a preliminary screening using the M.I.N.I., eligible couples completed a battery of questionnaires. These included the PCM (Hooley & Teasdale, 1989), the Beck Depression Inventory (Beck, Steer, & Brown, 1996),

and the Relationship Assessment Scale (Hendrick, 1988). For 21 subsequent days, each participant completed a diary questionnaire in private at the end of the day. Couples were instructed not to discuss their responses to the diary. Once a participant's responses for a given day were submitted, he or she could not return to them or edit them.

Independent Variables

Our primary independent variables consisted of global PC and 2 measures of criticism exposure: a partner-report of how much criticism is generally expressed towards the participant (i.e., global expressed criticism) and a participant-report of daily perceptions of their partner's criticality.

Global Perceived Criticism. Global PC was assessed using the PC measure (Hooley & Teasdale, 1989). Participants responded to the question "How critical is your partner of you?" on a 10-point Likert scale (Not at all – Extremely). We use the term "global PC" to explicitly distinguish it from a *daily* measure of perceived partner criticality that assesses criticism exposure on each day of the study. ("Global PC" in this study is the same as "PC" in other research using this measure.) The PC measure demonstrates good test-retest reliability (r = .75 over 5 months; Hooley & Teasdale, 1989).

Partner-Reported Expressed Criticism. The PC measure was also used to obtain a measurement of how much criticism respondents express towards their partners. Using the same 10-point Likert scale, participants responded to the question "How critical are you of your partner?" In our analyses, each participant's response to this question was used as a measure of his or her *partner's* exposure to criticism. Prior research indicates that spouses show moderate levels of agreement between PC and partner-reported criticism. For instance, Chambless and Blake (2009) reported a correlation of r = .43 between wives' PC ratings and husbands' ratings of much criticism they express; and a correlation of r = .39 between husband's PC ratings and wives' ratings much criticism they express.

Daily PC. To measure subjective criticism exposure each day, participants rated a single item: "My partner was judgmental, critical, or intolerant of me or others." Participants made ratings on 5-point Likert scales ranging from 1 to 5 (Not at all - Extremely). This operationalization allowed us to assess whether the participant was exposed to a generally critical or hostile environment that day. Though it does not guarantee that the participant was the target of criticism, our analyses suggest that responses to this question did reflect whether or not respondents felt personally criticized.

Covariates

PC is strongly correlated with relationship satisfaction (e.g., Chambless & Blake, 2009), and is sometimes (but not always; see Masland et al., 2015; Masland et al, *in press*) correlated with depression (Chambless & Blake, 2009; Peterson-Post, Rhoades, Stanley, & Markman, 2014; Smith & Peterson, 2008; White, Strong, & Chambless, 1998). Therefore, we included measures of these constructs so that we could rule out the possibility that PC's prediction of daily mood and cognition is due to the effects of depression or relationship satisfaction.

Depression. Participants completed the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996). The BDI-II is a 21-item measure of depressive symptomatology. On a four-point scale, respondents rate how much they have been bothered by each of 21 symptoms. Ratings are summed to provide a total depression score. The BDI is one of the most widely used measured of depressive symptomatology. It demonstrates high internal consistency (α = .93 among college students, α = .92 among outpatients; Beck et al., 1996) and construct validity. Alpha reliability in the current sample was good (α = .84).

Relationship Satisfaction. Participants completed the Relationship Assessment Scale (RAS; Hendrick, 1988). The RAS is a 7-item measure of relationship satisfaction. Participants answer each question on a 1 to 5 scale (e.g. 'How good is your relationship compared to most?'). Higher scores indicate greater relationship satisfaction. The RAS demonstrates convergent validity with measures of relationship satisfaction, such as the Dyadic Adjustment

Scale (r = .80; Hendrick, 1988). It demonstrates good reliability (α = 0.86) and agreement between spouses (r = .62) (Hendrick, 1988).

Dependent Variables

The daily diary was used to assess the following five moods and appraisals each day. Participants made all ratings on 5-point Likert scales ranging from 1 to 5 ("Not at all" to "Extremely [often]").

Depressed mood. Participants completed a modified version of the depression items of the Profile of Mood States (McNair & Lorr, 1964). Each day, participants rated the extent to which they felt "sad", "hopeless", and "discouraged," "irritable", "lonely", and "isolated". Responses for these 6 items were averaged together to form a single index of daily depressed mood (α = 0.84).

General well-being. Following a needs-satisfaction model of well-being (Deci & Ryan, 1995), we assessed several psychological needs thought to contribute to well-being, including appraisals of competence, autonomy, self-esteem, and life satisfaction (Pavot & Diener, 1993; Reis, Sheldon, Gable, Roscoe, & Ryan, 2000; Ryan & Deci, 2000). Participants rated their agreement with 13 statements addressing these constructs (e.g., "I felt as though I had the ability to solve my own problems", "I felt in control of my life", "I felt good about myself today", "I felt content with my life today."). Though these constructs are considered distinct, they were highly correlated in our data. For this reason, we combined them into a single composite measure indexing well-being related to the self. Internal reliability was excellent (Cronbach's α = 0.95).

Social connection. In addition to competence and autonomy, a feeling of social connection to others (often called *relatedness*) has also been considered a basic psychological need (Ryan & Deci, 2000). Participants rated their agreement with 6 statements, such as "I felt like I belonged or fit in with the people I spent time with today," and "I felt connected to people I interacted with today." Internal reliability was high (α = 0.90).

Perceived support. Perceived social support generally refers to the extent to which a person feels loved, cared for, and supported (Barrera, 1986). Low perceived support robustly predicts poor mental health outcomes, particularly for depression (Ibarra-Rovillard & Kuiper, 2011). Participants rated the extent to which they felt "loved", "valued", "accepted", and "supported" each day. Responses for these 4 items were averaged together (α = 0.90). Similar approaches have been used in diary studies to assess constructs such as partner closeness (e.g., Gleason, Iida, Shrout, & Bolger, 2008). Participants did not rate how much they felt supported by any particular person. Rather, they provided a global rating of how they felt each day.

Perceived partner understanding. Similar to the measurement of daily PC, participants rated the extent to which they perceived their partners to be understanding or sympathetic. Participants rated 5 statements, such as "When I explained things that happened to me today, my partner was able to understand these events from my perspective", "My partner was compassionate or sympathetic to others", and "My partner was sensitive other peoples' feelings." Internal reliability was high (α = 0.90).

Analytic strategy

All statistical analyses were conducted in R (R Core Team, 2014). To analyze daily outcomes, mixed-effects models were used to account for repeated measures, as well as dependencies caused by couples' relationships. In all models, we included a random intercept for each individual as well as each couple. We used the lme4 (Bates et al., 2014) and lmerTest (Kuznetsova et al., 2015) packages in R to conduct the mixed-effects analyses with significance testing.

To assess the shared and unique predictive validity of PC, we added variables to our models in multiple steps. For each dependent variable, we first tested a model including only PC and time (in units of days). Time was included in these models to account for the possibility that participants' responses to the diary questions were changing systematically over the course of

the study. For each dependent variable, we then conducted 3 subsequent models, with each model building on the previous one by introducing additional independent variables. In each model, all coefficients were estimated simultaneously. Thus, newly-added variables were allowed to reduce the coefficients of variables present in prior models. This approach allowed us to observe how the addition of variables changed the effect sizes of the pre-existing variables. In particular, we were interested in observing which covariates weakened the effects of global PC on our daily outcomes. With the exception of time, all variables were standardized using the mean and standard deviation of each variable, calculated with all available data.

Results

Descriptive statistics for all variables are provided in Table 5. Overall, as might be expected for a healthy and self-selected community sample, participants reported mid-range levels of PC, low levels of depression, and high relationship satisfaction. Interestingly, participants on average reported expressing more criticism than they received. Participants also reported low levels of daily depressed mood and low daily perceptions of partner criticality.

Table 5. Descriptive statistics for individual and daily variables

	n	Mean	SD	Possible Range
Person-level Variables				
Global PC	58	4.6	2.6	1 - 10
Expressed Criticism[1]	58	5.43	2.26	1 - 10
BDI	58	3.43	4.09	0 - 63
RAS	58	31.74	3.26	7 - 35
Daily-level Variables[2]				
Well being	1147	3.61	0.8	1 - 5
Depressed Mood	1147	1.48	0.58	1 - 5
Perceived Support	1147	3.88	0.78	1 - 5
Social Connection	1147	3.65	0.9	1 - 5
Partner Understanding	1018	3.46	1.06	1 - 5
Daily PC	1018	1.51	0.87	1 - 5

1. As reported by romantic partners
2. Descriptive statistics are calculated using all available data from eligible participants. Because the vast majority of participants completed between 18-21 daily responses, these statistics are not unreasonably biased by those participants who completed more responses.

Correlations between our independent variables are provided in Table 6. Global PC was significantly correlated with daily PC, partner-reported expressed criticism, and depression symptoms. Surprisingly, global PC was not correlated with relationship satisfaction, though relationship satisfaction was correlated with daily PC and partner-reported expressed criticism. Though modest, the convergence between daily PC and global PC, as well as between daily PC and daily perceived partner understanding, suggest that our operationalization of daily PC is indeed measuring daily experiences of criticism.

Table 6. Correlations among the primary independent variables

	Global PC		Daily PC		Daily Understanding		Expressed Criticism		BDI
Global PC									
Daily PC	**0.31**	*							
Daily Understanding	**-0.50**	***	**-0.45**	***					
Expressed Criticism[1]	**0.42**	***	0.23	†	**-0.46**	***			
BDI	**0.33**	*	**0.43**	***	**-0.35**	**	0.16		
RAS	-0.16		**-0.48**	***	**0.52**	***	**-0.33**	*	-0.2

1. As reported by romantic partners

†$p<.10$, *$p<.05$, **$p<.01$, ***$p<.001$

Mixed-effects analyses

Table 7 provides the results from our mixed-effects regression analyses. The first column (Model 1) describes the relationship between global PC and each of our dependent variables while controlling for time (in days) since couples started the diary portion of the study (range 1-21). Global PC predicted less positive daily well-being appraisals, $t(49.55) = -2.77$, $p = 0.008$; lower daily perceived support, $t(40.67) = -2.77$, $p = 0.008$; reduced daily social connection, $t(50.39) = -3.19$, $p = 0.002$; and lower daily perceived partner understanding,

Table 7. Standardized regression coefficients predicting daily well-being appraisals, depressed mood, perceived partner understanding, perceived support, and relatedness.

Variable	Model 1 ß	Model 1 SE	Model 2 ß	Model 2 SE	Model 3 ß	Model 3 SE	Model 4 ß	Model 4 SE
Well-being								
Time (in days)	**-0.01***	0.00	-0.01*	0.00	**-0.01***	0.00	**-0.01***	0.00
Global PC	**-0.25****	0.09	**-0.24****	0.09	**-0.24****	0.08	-0.15†	0.08
Daily PC	-	-	**-0.08*****	0.02	**-0.08*****	0.02	**-0.08*****	0.02
Expressed Criticism[1]	-	-	-0.12	0.09	-0.08	0.09	-0.08	0.08
RAS	-	-	-	-	**0.25***	0.09	**0.26****	0.09
BDI	-	-	-	-	-	-	**-0.27****	0.09
Depressed Mood								
Time (in days)	**-0.01*****	0.00	**-0.02*****	0.00	**-0.02*****	0.00	**-0.02*****	0.00
Global PC	0.05	0.09	0.03	0.09	0.03	0.09	-0.04	0.09
Daily PC	-	-	**0.14*****	0.03	**0.14*****	0.03	**0.14*****	0.03
Expressed Criticism	-	-	0.02	0.09	-0.01	0.09	-0.01	0.09
RAS	-	-	-	-	-0.15	0.09	-0.13	0.09
BDI	-	-	-	-	-	-	**0.28****	0.09
Partner Understanding								
Time (in days)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Global PC	**-0.35*****	0.09	**-0.30****	0.09	**-0.29*****	0.08	**-0.27****	0.09
Daily PC	-	-	**-0.11*****	0.02	**-0.11*****	0.02	**-0.11*****	0.02
Expressed Criticism	-	-	**-0.20***	0.09	-0.12	0.08	-0.13	0.08
RAS	-	-	-	-	**0.31*****	0.09	**0.31*****	0.09
BDI	-	-	-	-	-	-	-0.08	0.09
Perceived Support								
Time (in days)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Global PC	**-0.20****	0.07	**-0.22****	0.07	**-0.21****	0.06	**-0.17***	0.06
Daily PC	-	-	**-0.11*****	0.02	**-0.10*****	0.02	**-0.10*****	0.02
Expressed Criticism	-	-	**-0.17***	0.07	**-0.13***	0.06	**-0.13***	0.06
RAS	-	-	-	-	**0.32*****	0.07	**0.33*****	0.07
BDI	-	-	-	-	-	-	**-0.14***	0.07
Social Connection								
Time (in days)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Global PC	**-0.28****	0.09	**-0.27****	0.09	**-0.26****	0.08	**-0.20***	0.08
Daily PC	-	-	**-0.11*****	0.02	**-0.11*****	0.02	**-0.11*****	0.02
Expressed Criticism	-	-	-0.17†	0.09	-0.11	0.08	-0.11	0.08
RAS	-	-	-	-	**0.30*****	0.09	**0.30*****	0.08
BDI	-	-	-	-	-	-	**-0.17***	0.08

1. Expressed criticism refers to partner-reported belief as to how globally critical they are of the participant.

†$p<.10$, *$p<.05$, **$p<.01$, ***$p<.001$

t(54.6) = -3.9, *p* < .001. Global PC did not predict daily depressed mood, *t*(55.24) = 0.55, *p* = 0.58. (Here, *t*-test statistics for each regression coefficient refer to the null hypothesis test that the coefficient is equal to 0.)

The Model 2 column describes the relationships between all three criticism variables (global PC, daily PC, and partner-reported criticism) and each dependent variable. Here it is important to take note not only of which criticism variables are significant predictors of daily mood and appraisals, but also of how the regression coefficients for global PC change between Model 1 and Model 2 as a result of including the two "exposure" variables. Reduction in the regression coefficients of global PC would suggest that global PC's prediction of daily appraisals can be explained (either partially or fully) by exposure to criticism. To be clear, our two exposure variables include partner-reported criticism (Expressed Criticism in Table 7), which refers to each participant's partner's belief of how critical they are of the participant, globally; and daily PC, which refers to the participants' daily perceptions of how critical their partner was each day.

In general, global PC and daily PC were consistently linked with daily moods and appraisals, whereas partner-reported criticism was not. Daily PC was related to all five daily outcomes: less positive well-being, *t*(976.63) = -3.98, *p* < .001; greater depressed mood, *t*(986.57) = 5.44, *p* < .001; reduced perceived support, *t*(986.75) = -4.88, *p* < .001; reduced social connection, *t*(978.42) = -5.56, *p* < .001; and reduced perceived partner understanding *t*(979.21) = -5.17, *p* < .001. Partner-reported expressed criticism predicted only reduced daily perceived support, *t*(986.75) = -4.88, *p* < .001. Importantly, inclusion of these two criticism "exposure" variables did not appreciably weaken or change the relationship between global PC and daily moods and appraisals. Global PC remained a significant predictor of all daily variables except depressed mood. In other words, daily exposure to a critical partner does not explain global PC's prediction of daily well-being and social appraisals.

The primary purpose of Models 3 and 4 is to evaluate how relationship satisfaction and depression—two correlates of global PC and, thus, potential confounders—affect the

relationship between global PC and daily appraisals. Examining the Model 3 column, relationship satisfaction predicted greater well-being appraisals, $t(53.73) = 2.66$, $p = 0.01$; perceived partner understanding, $t(53.14) = 3.67$, $p < .001$; perceived support, $t(50.07) = 4.49$, $p < .001$; and social connection, $t(53.83) = 3.55$, $p < .001$; but not negative mood, $t(51.67) = -1.55$, $p = 0.13$. Similar to findings from Model 2, the inclusion of relationship satisfaction did not appreciably weaken or change the relationship between global PC and daily appraisals. It suggests that global PC's prediction of these daily variables is not merely an effect of relationship satisfaction.

Examining the Model 4 column, depression predicted 4 daily outcomes: less positive well-being appraisals, $t(53.1) = -3.06$, $p < .001$; greater depressed mood, $t(51.7) = 3.21$, $p < .001$; reduced perceived support, $t(50.34) = -2.07$, $p < .05$; reduced social connection, $t(53.64) = -2.1$, $p < .05$. Depression did not predict perceived partner understanding $t(53.55) = -0.92$, $p = 0.36$. The inclusion of BDI substantially reduced the effect of PC on well-being appraisals, rendering global PC a non-significant predictor of daily well-being appraisals. The inclusion of BDI also weakened global PC's relationship with the 3 socially-focused variables (social connection, perceived support, and perceived partner understanding), but these relationships remained statistically significant.

In summary, after controlling for daily criticism exposure, partner-reported expressed criticism, relationship satisfaction, and depression—global PC continued to predict all three daily socially-focused appraisals. Daily PC also remained a significant predictor of all criterion variables, though it consistently demonstrated the smallest effect sizes compared to PC, RAS, and BDI.

Discussion

The results of this study suggest that global PC ratings do not predict daily aspects of well-being simply because they provide a convenient index of exposure to criticism. When controlling for daily PC, as well as partner-reported expressed criticism, global PC ratings

remained a significant predictor of all dependent variables except depressed mood. Furthermore, even though global PC and daily PC ratings were correlated, controlling for daily PC negligibly reduced the effect sizes of global PC from Model 1 to Model 2. If global PC were merely a reflection of daily exposure to criticism, a more substantial reduction in these effect sizes should have been observed. Thus, our analyses suggest that PC may be related to daily cognition and appraisals through an independent process that does not hinge on daily exposure to criticism.

What else might explain global PC's prediction of daily social and well-being appraisals? Our analyses suggest that depression might partially explain these relationships. Inclusion of depression in our models reduced all effect sizes for global PC. In fact, depression completely accounted for the relationship between global PC and general well-being appraisals, rendering it non-significant. However, global PC continued to account for a significant amount of variance in all three socially-focused appraisals. Two of these variables, social connection and perceived support, tap into how people perceive and feel about their overall social environments—not merely their relationships with romantic partners. With respect to these variables, we found that people who report higher global PC tend to feel less connected to and supported by others each day. Thus, PC reveals how a person tends to feel each day about his or her overall social environment – not merely about the target relationship. This is consistent with research by Felix, Masland, and Hooley (2017) that indicates that PC ratings for a target-relationship are correlated with several characteristics and qualities of other relationships. Such similarities across a person's relationships, as well as PC's ability to tap into these similarities, supports the notion that PC is not merely a consequence or assessment of criticism exposure from the target person; rather, PC may reveal how a person generally feels about his or her social relationships—in other words, how connected, supported, and accepted he or she tends to feel each day.

It is also possible that PC is more than a passive indicator of social health: PC may be linked to a pervasive or trait-like negative cognitive bias that shapes how people experience their social world. Prior research has demonstrated that people who report higher PC also make more negative interpretations of ambiguous information and have difficulty disengaging attention from negative social information (Masland et al., 2015). This suggests that PC is not merely an indicator of attitudes, but perhaps indexes a process that actively constrains how people experience and evaluate their social interactions. Such a process would explain why people tend to report similar levels of criticism in many relationships. It might also explain why global PC is related to daily appraisals more strongly than daily partner criticism: PC may color all of a person's daily social experiences – not merely his or her interactions with a romantic partner. Thus, a person reporting high PC in one relationship may have a general bias towards perceiving criticism from many people each day.

These findings have implications for why PC predicts or contributes to clinical outcomes. From a prediction standpoint, PC behaves somewhat like a thermometer for basic psychological needs; it reveals a person's baseline tendency to feel competent, autonomous, supported, and socially connected each day. A substantial portion of well-being and psychological health can be explained by the extent to which these basic needs are met (Deci & Ryan, 1995; Reis et al., 2000; Ryan & Deci, 2000). Satisfaction of *social* needs may be particularly important. An extensive body of research indicates how feelings of disconnection, isolation, and loneliness can significantly harm mental health (for review, see Ernst & Cacioppo, 1999). Therefore, by providing insight into a person's daily experiences with respect to these needs, PC acts as a convenient indicator for a person's vulnerability to poor mental health outcomes.

Limitations

There are several limitations of this study. First, these data are correlational. Though we can make claims about PC's prediction of daily mood and appraisals, we cannot conclude that PC directly contributes to these moods and appraisals. Second, our operationalization of daily

PC may be too inclusive, insofar as it assesses general criticality of the partner, regardless of whether criticism was targeted at the respondent or a third party. This may have weakened the relationship between daily PC and other daily moods and cognitions. However, several of our analyses suggest that this is a valid operationalization of daily PC. As might be expected, daily PC was correlated with global PC, daily perceived partner understanding, depression, and relationship satisfaction. Furthermore, in regression analyses, daily PC was related to reduced well-being, more negative social appraisals, and greater depressed mood each day, even when all of its correlates were statistically controlled in Model 4. It is also worth noting that daily perceived partner understanding was measured similarly to daily PC: Of the five items used to assess daily partner understanding, three items allowed for partners' behavior to be targeted at others (e.g. "My partner was understanding of other people's beliefs, opinions, or points of view (including my own)"). If the relatively weak relationship between daily PC and daily well-being variables were merely an artifact of measurement, we would expect daily partner understanding to display similarly weak relationships. This was not the case, as daily partner understanding was strongly related to global PC ($r = -.50$), daily social connection, and daily perceived support.

Lastly, it is unclear if these findings generalize to clinical samples. It is possible, for instance, that in a clinical sample, daily exposure to criticism could have a larger effect on daily mood and appraisals. Clinical populations are more vulnerable and may have fewer emotional and cognitive resources to cope with and reduce the impact of partner criticism on daily well-being. Future studies should examine how global PC, daily criticism, and well-being are related in clinical populations.

Summary and Clinical Implications

The present study suggests that global PC ratings (i.e., those obtained using the original PC measure) do not predict clinical outcomes merely because they indicate a person's daily exposure to criticism. PC ratings also reveal diverse information about the extent to which a person's daily psychological needs are satisfied. Furthermore, PC ratings may reveal *trans-*

relational biases in how a person tends to experience and evaluate social relationships – not merely the target relationship.

These findings have may have clinical implications for the mitigation of adverse outcomes associated with PC. Because PC predicts daily well-being outcomes at least partially independently of daily criticism, interventions that focus on reducing expressed criticism among couples may address only part of the problem. Interventions must also target how one perceives and feels about his or her overall social world, for instance by boosting patients' feelings of social support and connection. This could include behavioral components, such as encouraging patients to develop new relationships or to invest more time in existing relationships that are perceived as safe and supportive. It may also be necessary for interventions to include cognitive components that address biased ways of perceiving and interpreting social experiences. PC has been linked not only with attention and interpretation biases (Masland et al., 2015), but also with criticality bias (Smith & Peterson, 2008), the over perception of criticism in interactions when none was intended or expressed. Cognitive interventions could focus on increasing awareness of and challenging dysfunctional thought patterns, such as mind reading, jumping to conclusions, or all-or-nothing thinking (Hooley & Miklowitz, 2017). Future research can refine these suggested interventions by identifying with greater specificity the processes that connect global PC ratings to daily needs-satisfaction, particularly social needs.

Paper 3: Perceived criticism is not merely a negative memory bias

Steven A. M. Felix & Jill M. Hooley

Manuscript in preparation

Abstract

Background: Perceived criticism (PC) refers to extent to which a person believes that his or her relative is globally critical of him or her. Previous research has implicated biased attention for negative information among people reporting high PC. It is believed that these attention biases can enhance encoding and subsequent recall for negative material, thereby increasing vulnerability to depression and other mental disorders. The present study investigates whether high PC is also characterized by enhanced recall for negative information.

Methods: Fifty-nine community participants were recruited from a university-based research participant pool. Participants completed a PC questionnaire, the Beck Depression Inventory, and the Self-Referent Encoding Task (SRET), an incidental recall task that requires self-referential processing of words with depressive themes.

Results: PC was unrelated to recall for positive or negative words, indicating the absence of a memory bias for words with depressive themes. Consistent with prior studies, however, depression scores were related to greater proportion of negative, compared to positive, recalled words.

Conclusions: These results suggest that PC is not characterized by biased encoding or recall of negative verbal information when it is processed self-referentially. Though this suggests that memory biases may not mediate the relationship between PC and poor clinical outcomes, future research must examine whether memory biases exist for different types of information, thematic content, and encoding methods.

Introduction

Perceived criticism (PC) refers to the amount of criticism a person perceives from a close relative (Hooley & Teasdale, 1989). A clinical construct, PC was originally proposed as a more practical way of assessing relationship criticism, which had typically been assessed using time-intensive standardized interviews, such as the Camberwell Family Interview (CFI; Leff & Vaughn, 1985). Though relationship criticism is highly related to clinical outcomes and could therefore provide useful information to clinicians (for recent review see Masland & Hooley, 2015), the time and training costs associated with interview methods make assessment of relationship criticism prohibitive for routine clinical use. In contrast, PC takes only a few minutes to administer and can predict clinical outcomes incrementally over objective interview methods (Hooley & Teasdale, 1989). This incremental prediction raises important questions about what extra information PC is providing over objective measures of criticism.

One possibility is that PC indexes clinically-relevant characteristics of the person making the rating, such as information processing biases. Using a modified version of the Attention Network Task (Fan, McCandliss, Sommer, Raz, & Posner, 2002) that employed angry and neutral faces as flankers, Masland, Hooley, Tully, Dearing, and Gotlib (2015) found evidence that PC was associated with attention bias. Specifically, individuals who reported higher PC had greater difficulty shifting attention away from negative pictures. However, PC was not associated with differences in speed of *orienting* to negative pictures. This pattern of results mimics findings related to attention bias in depression: Depressed and remitted-depressed individuals, as well as children of depressed mothers do not show enhanced orienting to negative pictures; but once they do attend to negative pictures, particularly those with depressive themes, they are more engaged with and distracted by them (Gotlib, Krasnoperova, Neubauer Yue, & Joormann, 2004; Joormann & Gotlib, 2007; Joormann, Talbot, et al., 2007). It is important to note that the attentional bias found in PC was not explained by depression, as PC was unrelated to depression scores on the BDI in the Masland et al. (2015) study.

58

Furthermore, because this attentional bias is found in children of depressed mothers, it may represent a pre-existing vulnerability to, rather than a consequence of, depression. This means that the attentional bias associated with PC may contribute to poor clinical outcomes in the same way that this bias contributes to depression.

Current thinking is that attentional bias contributes to the development and maintenance of depression because it fosters greater elaboration of depressogenic material and may therefore make it difficult to enact appropriate emotion-regulation strategies (Gotlib & Joormann, 2010). This can not only increase negative mood in the short term, but also enhance encoding and subsequent recall of negative information in the future. A number of cognitive models of depression explicitly posit such a relationship between attentional biases and memory biases (Joormann, Yoon, et al., 2007; Williams et al., 1997). For example, the framework proposed by Joormann and colleagues (2007) hypothesizes that deficits in cognitive control lead to difficulty disengaging attention from negative material. This difficulty disengaging leads to added elaboration and encoding, which in turn leads to memory biases.

Biased recall for negative information has been robustly demonstrated in depression, particularly on free recall tasks testing explicit memory (for review see Gotlib & Joormann, 2010). Memory biases are particularly pronounced when the incidental learning component entails semantic or self-referent processing of the stimuli (e.g., words; Derry & Kuiper, 1981). This may be due to the tendency of depressed individuals, as well as those at risk for depression, to ruminate and elaborate more extensively on self-relevant information. Gotlib and colleagues (2004) demonstrated a memory bias in depressed and socially anxious groups using the Self-Referential Encoding Task (SRET). In this task, participants are presented with positive and negative words and asked to decide whether or not each word describes them. After rating all of the words, participants engage in a brief distractor task and are then asked to recall as many of the words as possible, regardless of their endorsement of each word. The researchers found that relative to non-psychiatric controls, the depressed and socially anxious groups

59

recalled more negative words (regardless of self-endorsement). Furthermore, among the words that were endorsed as self-referent and then subsequently recalled, for the depressed group, a greater proportion of these words were negative.

Present study

The theoretical and empirical links between attention and memory biases in depression suggest that memory biases may also be associated with PC. In the present study we investigated whether people who report higher PC demonstrate biased recall of negative information. An understanding of whether this bias presents in those reporting high PC would allow us to further characterize the cognitive factors that contribute to PC and also enhance understanding of the relationship between PC and clinical outcomes. The cognitive biases associated with PC, such as attention and interpretation biases (Masland et al., 2015), are also implicated in the onset and maintenance of depression. This overlap in cognitive biases suggests one potential route by which PC could be related to depression and other clinical outcomes. Presence of a memory bias could also provide some insight into whether PC ratings themselves may be impacted by a negative memory bias.

In this study we evaluated whether PC was associated with stronger encoding and subsequent recall of negative words when these words are processed self-referentially. Given that individuals reporting high PC demonstrate difficulty disengaging their attention from negative material, we hypothesized that higher PC ratings would be associated with enhanced recall for negative words. We tested this using the Self-Referent Encoding Task (SRET; Derry & Kuiper, 1981), which has been widely used in the clinical literature to study memory bias.

Methods

Participants

61 participants were recruited from the Harvard University Psychology Department's "Study Pool" system. All participants were undergraduates at Harvard or other local colleges. Because of the semantic demands of the task used in this study, participants were also required

to be fluent English speakers. One participant was excluded because of an error during the study procedure that invalidated her data, and a second participant was excluded due to unusual task performance on the SRET. The final sample consisted of 59 participants.

Measures

Perceived Criticism Measure (PCM; Hooley & Teasdale, 1989). The PCM is a single-item self-report scale that measures criticism in a person's closest relationship. Participants first select their most important relationship and then respond to the question "How critical is [person] of you?" The rating is made on a 10-point Likert scale ranging from "Not at all critical" to "Extremely critical". In this study, the PCM was used to assess PC not only from a person's most important relative, but also from up to 5 additional people, including romantic partners, mothers (or mother-figures), fathers (or father-figures), best friends, and a 5th person of each participant's choosing. By collecting multiple PC ratings, we could calculate an average PC rating to represent a person's general level of PC across their most important relationships (Felix, Masland, & Hooley, 2017). The PCM demonstrates convergent validity with trained and untrained raters' observations of criticism (Chambless & Blake, 2009; Gerlsma et al., 2009; Hooley & Teasdale, 1989) and is reliable when respondents rate their most important relationships (5-month test-rest reliability r = .75; Hooley & Teasdale, 1989). The PCM also demonstrates discriminant validity with respect to mood, neuroticism, and other maladaptive personality traits (Masland, Drabu, & Hooley, in press; Riso, Klein, Anderson, Ouimette, & Lizardi, 1996). Though PC is sometimes correlated with measures of depressive symptomatology and global functioning (e.g., Chambless & Blake, 2009; Chambless et al., 2001), these correlations are modest and inconsistent, indicating that PC is not likely a proxy for psychopathology or emotional distress.

Beck Depression-Inventory-II (BDI-II; Beck et al., 1996). Because memory biases are associated with depression, we controlled for depression symptomatology using the BDI-II. This is a 21-item self-report measure to that asks respondents report how much they have been

bothered by various symptoms using a four-point scale. The BDI is one of the most widely used measures of depressive symptomology. It demonstrates high internal consistency (α = .93 among college students, α = .92 among outpatients; Beck et al., 1996). The BDI-II reliably distinguishes between depressed and non-depressed individuals, though it often correlates with measures of self-reported anxiety (Richter, Werner, Heerlein, Kraus, & Sauer, 1998).

Self-Referent Encoding Task (SRET; Derry & Kuiper, 1981). The SRET is a brief, computer-based memory task that has been use to demonstrate negative memory biases in depression (e.g., Gotlib, Kasch, et al., 2004). It consists of three parts: an incidental learning (i.e., encoding) component, a brief distractor task, and then a free recall period. During the incidental learning component, participants are seated at a computer screen and are instructed to focus on a target in the middle of the screen. In each trial, the target is replaced by the words "Describes me?" for 500ms. Next there is a 250-ms pause, after which a positive or negative adjective is presented. Participants then respond "yes" or "no" using two computer keys. Participants are instructed to respond as fast as they can. There is a 1000-ms pause between trials. There are 20 trials involving negative adjectives and 20 trials involving positive adjectives. We used a version of the task that was identical to the one used by Kircanski, Mazur, and Gotlib (2013) and Asarnow, Thompson, Joormann, and Gotlib (2014), who derived the adjectives from studies of cognitive biases in depression (e.g., Gotlib, Kasch, et al., 2004). Immediately following the encoding portion of this task, participants completed the Digit Substitution Test (described below) as a distractor task lasting a total of 3 minutes (1 minute for reading instructions and 2 minutes for the task). Afterwards, participants were instructed to think back to the incidental learning component and write down as many words as they could remember, ignoring whether or not they felt the words described them.

We computed 3 indices for memory bias. To measure the presence of a general *recall bias* towards remembering negative words, we calculated the proportion of all recalled words that were negative. This score ignores whether or not words were endorsed as self-referent by a

participant. Thus, a person who recalled 4 negative and 6 positive words would have a recall

bias score of 0.4. A score significantly higher than 0.5 indicates a negative recall bias, whereas

a score significantly lower than 0.5 indicates a positive recall bias. This index is significance-

tested for positive or negative words (but not both), as each yields the same result only with

opposite signs.

We also calculated a score that indexed the balance between positive and negative

words among all self-referent recalled words. It is calculated by dividing the number of negative

or positive self-refer   ent recalled words by the *total* number of self-referent recalled words (i.e.,

positive plus negative). Again, this index need only be calculated and significance-tested for

positive or negative words, as the two values sum to 1. For instance, if a person recalled 2

negative self-referent words and 6 positive self-referent words, his score would be 0.25 (or .75 if

the number of positive self-referent recalled words is used in the numerator). Though this

proportion has been used in prior studies to index memory bias (Fritzsche et al., 2010; Gotlib,

Kasch, et al., 2004; Kircanski et al., 2013), it confounds recall bias with endorsement bias (e.g.,

the tendency to endorse more positive words than negative words), which is present in each of

the aforementioned studies and also covaries with depression. To maintain comparability with

past studies, we decided to include this index but also calculate a new index that separates

endorsement from recall.

To calculate a pure *self-referent recall bias* we simply normalized the aforementioned

index (separately for both positive and negative words) by dividing it by a person's rate of

negative or positive endorsement. For instance, in the case of negative self-referent recall bias,

if the person described above had endorsed 4 negative and 12 positive words, we would divide

the score of 0.25 at the first step by the negative endorsement rate of 0.25 (4/16), leading to a

final score of 1. A score of 1 indicates no self-referent recall bias because it suggests that

among the pool of words deemed to be self-referent, a person's probability of recalling a

negative word from this pool is the same as the proportion of negative words available for

selection from pool. In the case of the negative index, proportions greater than 1 indicate biased recall for negative self-referent words, and proportions lower than 1 indicate biased recall for positive words. The opposite is the case for the positive index.

Wechsler Adult Intelligence Scale (WAIS) – Digit-Symbol Substitution Test (DSST; Wechsler, 1981). The DSST requires participants to translate digits (0-9) to symbols using a key. Participants have 2 minutes to translate as many digits as possible. In this study, the DSST served as a distractor task between the incidental learning and free recall portions of the SRET. Performance on the DSST was not officially scored, but response sheets were examined to ensure that participants were focused on the task during the 2-minute period.

<div align="center">

Results

</div>

Sample-wide SRET Performance

Descriptive statistics are displayed in Table 8. Dependent-samples t-tests indicated that, as a whole, participants both endorsed, $t(58) = 10.23$, $p<.001$, $d = 1.33$, and recalled $t(58) = 5.01$, $p<.001$, $d = 0.65$, more positive words than negative words. Furthermore, a one-sample t-test indicated that the overall recall bias index was significantly different from 0.5, $t(58) = -4.76$, $p<.001$, $d = -0.62$. This indicates a sample-wide average recall bias for positive words when ignoring whether or not participants endorsed particular words.

Next we examined the proportions of positive versus negative recalled self-referent words, which is the index used in prior studies to assess self-referent recall bias. Using a one-sample t-test to examine whether this value is significantly different from 0.5, we found that a significantly greater proportion of self-referent recalled words were positive $t(58) = 10.22$, $p<.001$, $d = 1.33$. However, as described in the Measures section, we believe this index is confounded by positive versus negative endorsement rates, and cannot be used as an index of recall bias. Indeed, the correlation between this index and endorsement rates for positive words is very large, $r = .82$, $p<.001$; and, in Table 8, the proportions of negative and positive endorsed words are nearly identical to the proportions of negative and positive recalled endorsed words.

<div align="center">

64

</div>

Table 8. Descriptive statistics for the self-report measures and SRET performance

	n	M	SD	Min - Max	n	M	SD	Min - Max
Self-report Measures								
Age	56	21.8	3.18	18.33 - 39.67				
BDI	59	7.14	7.59	0.00 - 31.00				
PC - Average	59	4.07	1.67	1.00 - 8.80				
PC - Primary Relative	59	4.03	2.61	1.00 - 10.00				
SRET Performance		**Negative words**				**Positive words**		
Number of endorsed words	59	5.63	4.09	0.00 - 16.00	59	15.27	4.02	3.00 - 20.00
Number of recalled words	59	6.1	2.07	2.00 - 10.00	59	8.05	2.4	3.00 - 13.00
Proportion of endorsed words by valence[1]	59	0.26	0.19	0.00 - 0.80	59	0.74	0.19	0.20 - 1.00
Proportion of recalled words by valence[2]	59	0.43	0.11	0.13 - 0.70	59	0.57	0.11	0.30 - 0.87
Proportion of recalled endorsed words by valence[3]	59	0.24	0.19	0.00 - 0.75	59	0.76	0.19	0.25 - 1.00
Self-referent recall bias[4]	58	1.02	0.86	0.00 - 4.75	59	1.05	0.19	0.59 - 1.73

1. Can be considered an index for positive/negative endorsement bias. Proportions sum to 1.
2. Can be considered an index for positive/negative recall bias. Proportions sum to 1.
3. Index of self-referent recall bias used in prior studies. Proportions sum to 1.
4. Novel index of self-referent recall bias. Bias is indicated by values close to 1. The ratio of Proportions do not sum to 1.

Thus, when this index is normalized by each person's rate of positive versus negative endorsement, there is no evidence of a pure self-referent recall bias. Both one-sample t-tests examining whether positive and negative self-referent recall bias indices were significantly different from 1 were non-significant, though there was a slight trend for participants to recall more positive endorsed words, $t(58) = 1.81$, $p = 0.08$, and $t(58) = 0.17$, $p=0.87$, respectively. This means that, on average, participants recalled positive and negative self-referent words in the same proportions as they were endorsed. For instance, if 30% of a person's endorsed words were negative, then approximately 30% of this person's recalled self-endorsed words were also negative.

Correlations

Correlations among our key variables are displayed in

Table *9*. Neither average PC nor PC from the primary relative was related to depression symptomatology. Furthermore, PC was not related to any performance index on the SRET. In contrast, higher BDI scores were associated with endorsing more negative words relative to positive words, recalling more negative words relative to positive words (i.e., negative recall bias), and recalling more negative self-referent words relative to positive self-referent words. BDI was not associated with any self-referent recall bias. In order to illustrate the relationships between BDI, PC, and recall bias, we provide scatterplots with regression lines in Figure 1. At the highest levels of BDI, more than 50% of recalled words are negative, whereas at the lowest levels, much less than 50% of recalled words are negative. No such relationship appears for PC.

Table 9. Correlations among self-report measures and SRET performance

	BDI	PC - Average	PC - Primary Relative
BDI			
PC - Average	-0.054		
PC - Primary Relative	-0.185	0.784***	
Negative proportion of endorsed words[1]	0.666***	-0.009	-0.19
Negative proportion of recalled words[1]	0.348**	-0.074	-0.215
Negative proportion of recalled self-referent words[1]	0.574***	-0.026	-0.223
Negative self-referent recall bias	-0.133	-0.17	-0.166

p<.05, **p*<.01, ***p*<.001

1. The negative and positive proportions add to 1, the correlations have the same magnitude but opposite signs.

Figure 1. Scatter plots and regression lines for recall bias according to BDI and PC Average
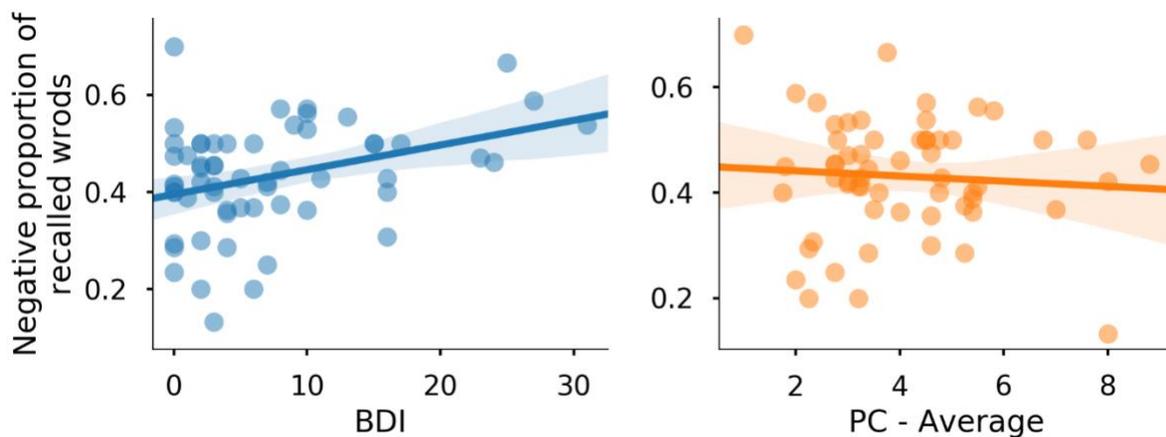
Discussion

We found no evidence to suggest that PC was related to biased memory for negative information, even when this information is processed self-referentially. We used an incidental-recall task that has been used to demonstrate the presence of memory bias in depression (Gotlib, Kasch, et al., 2004). Consistent with this research, we found that depression scores on the BDI were related to greater endorsement of negative words, a higher ratio of negative to positive recalled self-referent words, and a higher ratio of negative to positive recalled words (ignoring word endorsement). Neither a person's average PC rating nor PC rating for a primary relative was correlated with any of these indices. That task performance was related to depression as expected indicates that these null relationships are unlikely to be explained by an error in the task or procedure.

The absence of a relationship between PC and *general* recall bias (i.e., examining all words, not only endorsed words) is surprising in light of prior research demonstrating attentional bias among people reporting high PC. Similar to findings from the depression literature (for review see Gotlib & Joormann, 2010), people reporting high PC have difficulty disengaging attention from negative visual information (Masland et al., 2015). Cognitive theories of depression posit a causal link between increased attention to negative information—and thus greater elaboration and memory encoding—to enhanced recall for negative information (Joormann, Yoon, et al., 2007). The discrepancy between Masland and colleagues' (2015) detection of an attention bias and our finding no evidence for a memory bias in PC can possibly be explained by differences in task stimuli. Masland and colleagues demonstrated enhanced attention to pictures of angry faces, whereas the task used in the present study employed word stimuli (e.g., awful, lonely, hopeful). It is possible that PC could be associated with memory bias for negative visual, but not verbal, information. Furthermore, there may also be specificity in the themes for which people with high PC demonstrate heightened attention. For example, PC may be associated with biased attention for social-evaluative information (e.g., angry faces), as has

68

been found in social phobia. Future research must use different types of stimuli and thematic content to probe the specific contexts in which PC is related to attention or memory biases.

The absence of a relationship between PC and our two self-referent word indices is less surprising. As discussed in the Methods section and substantiated in our results, the index for self-referent memory bias used in previous studies—the negative (or positive) proportion of self-referent recalled words—is highly related to the ratio of negative to positive endorsement. Thus, the lack of relationship between PC and this index can be explained by the fact that PC was similarly unrelated to positive or negative endorsement. That is, people reporting high PC did not view themselves as more negative or less positive than those reporting low PC. Similarly to PC, the significant relationship between BDI and the negative proportion of self-referent recalled words (i.e., the index found in prior studies) can be explained simply by greater endorsement of negative words among more depressed participants. The absence of a relationship between our novel index of self-referent recall bias and PC or BDI is similarly unsurprising. Though prior studies have not explicitly calculated and analyzed this index, some have provided the values (e.g., averages, proportions) necessary to do the computation (e.g., Gotlib, Kasch, et al., 2004). Our examination of these values indicates that they would be unrelated to BDI in these studies, which is consistent with the results of the present study.

The present findings have important implications for both our understanding of the PC construct as well as its relationship with clinical outcomes. Prior research has indicated that PC ratings to some extent reflect characteristics or biases of the individual, rather than objective criticism (Felix, Masland, & Hooley , 2017; Peterson et al., 2009; Smith & Peterson, 2008). With regards to the PC construct, the current findings suggest that higher PC ratings cannot be explained by a bias in the way people bring to mind memories about the relationship when making a PC rating. In other words, it is not the case that people who report high PC are simply better at recalling negative memories. Criticality bias—the tendency to perceive criticism in an interaction when none was expressed (Smith & Peterson, 2008)—may instead result from

69

negatively biased interpretations of events, which has been implicated in PC. For instance, Masland and colleagues (2015) found that when people listened to negative words digitally morphed with similar-sounding neutral words (e.g., sad and sand), those reporting high PC were more likely to report hearing the negative word. This kind of basic interpretation bias could foster criticality bias in daily interactions, which in turn could augment overall PC. Smith and Peterson (2008) have reported a strong correlation between criticality bias in laboratory interactions and global PC ratings.

These findings also suggest that memory bias may not contribute to the relationship between PC and poor clinical outcomes. It has been suggested that PC's predictive validity may partially be due to its ability to capture a wide variety of information about the individual, including cognitive vulnerabilities like dysfunctional attitudes or attentional biases, perceptions and quality of social relationships and support, and existing symptomatology (Felix, Hooley, & Hooker, 2017; Felix, Masland, & Hooley, 2017). Given that memory biases may have a causal role in the development and maintenance of depression, the present study suggests that memory biases for negative verbal stimuli may not contribute to or account for PC's relationship with depression.

The present study has a number of limitations. The participants in this study are a convenience sample of undergraduates at an elite institution. As such, they may not be representative of the general population, particularly with respect to recall ability. It is possible that having an above average ability to recall words could dilute or remove a bias that exists among people with average memory abilities. However, if this were the case, we might also expect BDI to be unrelated to recall bias. Thus, it is not likely that the lack of association between PC and SRET performance is an artifact of the sample. Another limitation of this study is that it employed only verbal stimuli. Studies of information processing biases in psychopathology have often found that mental disorders demonstrate specificity in the types of stimuli (e.g., verbal, visual) and thematic content (e.g., depressive/sad, social-evaluative, threat)

70

for which biases can be observed (Gotlib & Joormann, 2010; Mathews & MacLeod, 2005). Thus, it is possible that a memory bias may be observed in PC using other stimuli and thematic content. Given that prior research has found enhanced attention towards pictures of angry faces in PC (Masland et al., 2015), future research investigating memory bias in PC should utilize visual cues or social-evaluative stimuli.

Despite these limitations, the present research has several strengths. We employed a well-tested task that consistently elicits a negative recall bias in samples with depression, vulnerability to depression, and in some cases social phobia (Gotlib, Krasnoperova, et al., 2004). This gives us confidence that the lack of association between PC and recall bias for verbal, depressive stimuli is genuine. Second, we calculated a novel index for self-referent recall bias that effectively removes the impact of endorsement rates on the index used in previous studies. Examination of descriptive data provided in previous studies suggests that the lack of relationship between depression and pure self-referent recall bias found in this study would be consistent with past studies if they were to employ this new index.

General Discussion and Conclusion

Summary of Findings

PC has the potential to be a valuable clinical tool. Its value, however, is limited by ambiguities in the meaning of the construct. If we do not know exactly what PC is measuring, we cannot develop and test theory-driven hypotheses explaining its relationship with clinical outcomes. Furthermore, we cannot develop appropriate interventions that target the underlying processes that put patients at risk. The studies presented in this dissertation were designed to sketch a more holistic understanding of the PC construct. Specifically, these studies were designed to test the extent to which PC provides information beyond criticism and beyond the target relationship—for instance, information about the individual and his or her cognitive traits and broader social environment.

In study 1, we examined the extent to which PC should be conceptualized as an individual versus a relationship variable. We found significant evidence that PC may provide information in equal measure at both levels. Specifically, we found that individuals tend to report similar levels of criticism across all of their important relationships and that these PC ratings can be reliably averaged together to gain an overall sense of a person's baseline level of PC across many relationships. Furthermore, regression analyses indicated that a person's average PC in a set of relationships is the strongest of predictor of PC in another relationship, followed then by relationship-specific variables such as relationship satisfaction vis-à-vis romantic partners and parental attachment vis-à-vis parents. We also found that PC ratings, particularly average PC, were correlated with several clinical risk factors, including existing global symptomatology, dysfunctional attitudes, low perceived social support, and insecure attachment. Together these findings suggest that although experiences in a particular relationship have a strong influence on PC in that relationship, there is a strong individual contribution to a PC rating. This contribution might be cognitive (e.g., attention, interpretation biases; Masland et al., 2015), attributional (Allred & Chambless, 2014), or affective. Additionally, PC may predict clinical

72

outcomes because it indicates a broadly elevated risk-profile across a number of relevant clinical dimensions.

Study 2 examined whether global PC from a romantic partner predicts well-being outcomes because it indexes daily encounters with a critical partner. In other words, we tested whether daily PC fully or partially accounted for the relationship between global PC ratings and aspects of daily well-being. We found that although global PC and daily PC were correlated, global PC was a much stronger predictor of daily well-being outcomes than daily PC; furthermore, the inclusion of daily PC in hierarchical regression models failed to reduce the effect of global PC on all daily well-being outcomes. Depression symptomatology, on the other hand, significantly altered the predictive power of global PC, limiting its prediction to only daily *social* appraisals. Specifically, people reporting higher global PC from their partners also tended to feel less connected, understood, and supported by others each day. These findings suggest that PC provides unique information about how a person tends to view and appraise his social environment each day. PC may predict clinical outcomes because it not only shares variance with pre-existing symptomatology but also indexes this more general social vulnerability.

Based on cognitive theories of depression, study 3 tested the hypothesis that people reporting high PC would demonstrate enhanced recall for negative information when it is processed self-referentially. This hypothesis was not supported. Though depression was linked with negative memory bias, as expected, PC was not. These findings suggest that PC cannot be dismissed as a negative bias in the way a person recalls information about the relationship. Furthermore, they suggest that enhanced recall for negative self-referent material does not explain the relationship between PC and subsequent clinical outcomes.

Implications

Together, these studies indicate that PC cannot be conceptualized merely as an indicator of how much criticism a person objectively receives from a relative, nor how much criticism a person subjectively perceives. This does not mean that criticism plays no role in PC;

on the contrary, there is consistent evidence that PC is correlated with other measures of criticism, including spouses' ratings of how much criticism they express and untrained raters' impressions of spouses' criticism in a laboratory interactions (e.g., Chambless & Blake, 2009; Smith & Peterson, 2008). However, these correlations tend to be small to medium, suggesting that *objective* criticism may play only a minor role in PC. Furthermore, our findings from paper 2 (Felix, Hooley, & Hooker, 2017) provide an important extension to this conclusion: Even daily *perceived* criticism may play only a minor role in global PC ratings, as daily PC and global PC were only modestly correlated. Thus, small or null correlations between global PC and objective criticism measures found in the literature cannot be interpreted simply as the natural discrepancy between objective and subjective ratings. Instead, it seems that global PC ratings depend only modestly on criticism—whether objectively or subjectively assessed.

What, then, is PC? The studies presented here suggest that PC is a diverse construct that reflects many process and constructs at both the relationship and individual levels. Critically, it is not redundant with any one particular variable. It may be appropriate to conceptualize PC as an all-purpose risk factor that indicates a generally elevated profile of vulnerabilities, particularly in regard to social factors. Thus PC's predictive validity may not be derived from its ability to index any one factor but rather its ability to provide a small amount of information about a wide variety of clinically-relevant factors, including information processing biases, attributional biases, and quality of social relationships. In fact, what may make PC unique and predictive of clinical outcomes is its capacity to provide information about a person's social relationships beyond what can be captured by symptomatology or relationship satisfaction. Across the existing literature as well as the current studies, PC is reliably linked with social and relationship variables such as attachment, relationship quality, and relationship satisfaction. Furthermore, the research presented here demonstrates how PC—even when measured in a single relationship—uniquely predicts daily social evaluations, including perceptions of partner understanding and general feelings of social belonging and support.

Based on this finding in Study 2, we present the hypothesis that measuring PC may allow us to tap into a more general social *lens* or filter that shapes how a person perceives and experiences their social world. This might explain not only why PC is correlated with the quality of several social relationships, but also why it predicts clinical outcomes incrementally over symptomatology and relationship adjustment (Hooley & Teasdale, 1989).

Limitations and Future Directions

Though several limitations have been described in each of the presented studies, there are more general limitations and concerns with regard to the broader conclusions derived from this work. One critical question is whether the results of these studies can be generalized to clinical populations. In other words, does PC have the same meaning and correlational structure in clinical and non-clinical populations. Unfortunately, there is not a clear answer to this question, and the current literature provides few hints. Research on PC—the present studies being no exception—has tended to be conducted with either clinical or non-clinical samples, but rarely both. However, one study included both non-depressed couples and couples with one depressed partner to test the relationship between negative social attributions and PC (Peterson & Smith, 2011). The researchers found that negative attributions about spousal behavior accounted for a significant amount of variance in PC even after controlling for marital adjustment, measures of objective criticism, and diagnostic group. Importantly, not only was diagnostic group not related to PC, it also did not moderate the relationship between PC and other factors, suggesting that PC's links to the variables under investigation was the same in healthy and pathological samples. Future studies must take on a similar approach of including both clinical and non-clinical samples in the same study in order to investigate more thoroughly whether findings vary across populations.

Another key limitation to this research is the overall lack of racial diversity in the samples. There is reason to think that PC may develop differently across cultures due to variation in cultural relationship norms and ways of making attributions about relatives' behavior,

including criticism (Allred & Chambless, 2014). However, this problem also supports the overall notion that PC is not a straightforward assessment of expressed criticism, but rather an assessment of the unique (or cultural) ways a person perceives, interprets, and makes attributions about others' behavior. Future studies may extend the studies presented here by examining whether differences in cultural specific attributions or relationship expectations can fully account for any racial differences in PC and its links with other key variables.

A final consideration is if and how these findings can be incorporated into clinical practice. To date, no studies have been conducted that examine interventions among patients reporting high PC, neither to reduce PC itself or to mitigate the negative outcomes associated with it. This is unfortunate because the PC construct itself was created with the hope that it could help identify patients at high risk of relapse and thus provide an opportunity to supply these patients with appropriate interventions (Hooley & Richters, 1991; Hooley & Teasdale, 1989). Though many questions about PC remain, there is enough theoretical perspective on PC to generate hypotheses as to which types of interventions might be effective. Until researchers begin to design and test prospective interventions on PC or utilizing PC, the clinical potential of PC research will remain unrealized.

Although the findings presented here do not point to a single intervention target that would comprehensively reduce risk of poor clinical outcomes, there are a few general areas that can guide future ideas about clinical interventions. In demonstrating that PC reflects both individual and relationship processes, Study 1 suggests that interventions must target the patient just as much as they target his or her primary relationships. Study 2 builds upon this by highlighting social perceptions and evaluations as a primary target for intervention. Results of Study 2 suggest that people who report higher PC tend to feel less connected to and supported by others each day and also tend to perceive their partners as less understanding each day. Thus, interventions might target how a person perceives and feels about his or her social experiences, for instance by boosting feelings of social support and connection. This could

76

include behavioral components, such as encouraging patients to develop new relationships or to invest more time in existing relationships that are perceived as safe and supportive. This might also include cognitive components that address biased ways of perceiving and interpreting social experiences.

Conclusion

Considering the present research in the context of existing literature, a more nuanced understanding of the PC construct is emerging. This understanding embraces the complexity of PC as a measure of both relationship experiences and the internal processes that shape how a person contributes to and creates these experiences through personal ways of perceiving and evaluating relationship events—as well as through their own behaviors. This nuanced understanding also helps us understand why PC reliably predicts clinical and well-being outcomes over any number of related or confounding constructs: It provides a small amount of information about a diverse swath of clinically-relevant factors without being redundant with any one. Although researchers of PC must continue to enhance our understanding of PC, it is becoming increasingly important to bring to fruition Hooley and Teasdale's (1989) vision for PC to be an actionable clinical assessment. This will require future researchers to develop and test interventions targeted at reducing PC or the processes theorized to underlie PC's relationship with relapse and treatment outcome.

References

Allred, K. M., & Chambless, D. L. (2014). Attributions and race are critical: Perceived criticism in a sample of African American and white community participants. *Behavior Therapy*, *45*(6), 817–830. https://doi.org/10.1016/j.beth.2014.06.002

Andreu, Y., Galdón, M. J., Dura, E., Ferrando, M., Murgui, S., García, A., & Ibáñez, E. (2008). Psychometric properties of the brief symptoms inventory-18 (BSI-18) in a Spanish sample of outpatients with psychiatric disorders. *Psicothema*, *20*(4), 844–850.

Asarnow, L. D., Thompson, R. J., Joormann, J., & Gotlib, I. H. (2014). Children at risk for depression: Memory biases, self-schemas, and genotypic variation. *Journal of Affective Disorders*, *159*, 66–72. https://doi.org/10.1016/j.jad.2014.02.020

Bachmann, S., Bottmer, C., Jacob, S., & Schröder, J. (2006). Perceived criticism in schizophrenia: A comparison of instruments for the assessment of the patient's perspective and its relation to relatives' expressed emotion. *Psychiatry Research*, *142*(2–3), 167–175. https://doi.org/10.1016/j.psychres.2005.10.001

Baron, R. A. (1988). Negative effects of destructive criticism: Impact on conflict, self-efficacy, and task performance. *Journal of Applied Psychology*, *73*(2), 199–207. https://doi.org/10.1037/0021-9010.73.2.199

Barrera, M. (1986). Distinctions between social support concepts, measures, and models. *American Journal of Community Psychology*, *14*(4), 413–445. https://doi.org/10.1007/BF00922627

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., & Singmann, H. (2014). Package "lme4" version 1.1-6.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037//1089-2680.5.4.323

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper & Row.

Beck, A. T., Steer, R. A., & Brown, O. K. (1996). *Beck Depression Inventory manual (2nd ed.)*. San Antonio, TX: Psychological Corporation.

Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, *43*(October 2015), 800–813. https://doi.org/10.3758/s13428-011-0081-0

Bernstein, D. P., Ahluvalia, T., Pogge, D., & Handelsman, L. (1997). Validity of the Childhood Trauma Questionnaire in an adolescent psychiatric population. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(3), 340–348. https://doi.org/10.1097/00004583-199703000-00012

Bradbury, T. N., & Fincham, F. D. (1990). Attributions in marriage: review and critique. *Psychological Bulletin*, *107*(1), 3–33. https://doi.org/10.1037/0033-2909.107.1.3

Brown, G. W., Birley, J. L. T., & Wing, J. K. (1972). Influence of family life on the course of schizophrenic disorders: A replication. *The British Journal of Psychiatry*, *121*(3), 241–258. https://doi.org/10.1192/bjp.121.3.241

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Butzlaff, R. L., & Hooley, J. M. (1998). Expressed emotion and psychiatric relapse. *Archives of General Psychiatry*, *55*(6), 547. https://doi.org/10.1001/archpsyc.55.6.547

Chambless, D. L., & Blake, K. D. (2009). Construct Validity of the Perceived Criticism Measure. *Behavior Therapy*, *40*(2), 155–163. https://doi.org/10.1016/j.beth.2008.05.005

Chambless, D. L., Bryan, A. D., Aiken, L. S., Steketee, G., & Hooley, J. M. (2001). Predicting expressed emotion: a study with families of obsessive-compulsive and agoraphobic outpatients. *Journal of Family Psychology*, *15*(2), 225–240. https://doi.org/10.1037/0893-3200.15.2.225

Chambless, D. L., & Steketee, G. (1999). Expressed emotion and behavior therapy outcome: a prospective study with obsessive-compulsive and agoraphobic outpatients. *Journal of Consulting and Clinical Psychology*, *67*(5), 658–665. https://doi.org/10.1037/0022-006X.67.5.658

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, *98*(2), 310–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3901065

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS ONE*, *8*(3). https://doi.org/10.1371/journal.pone.0057410

Cutrona, C. E., & Russell, D. W. (1987). The provisions of social relationships and adaptation to stress. *Advances in Personal Relationships*, *1*, 37–67.

de Graaf, L. E., Roelofs, J., & Huibers, M. J. H. (2009). Measuring dysfunctional attitudes in the general population: The Dysfunctional Attitude Scale (Form A) revised. *Cognitive Therapy and Research*, *33*(4), 345–355. https://doi.org/10.1007/s10608-009-9229-y

Debrot, A., Cook, W. L., Perrez, M., & Horn, A. B. (2012). Deeds matter: daily enacted responsiveness and intimacy in couples' daily lives. *Journal of Family Psychology : JFP : Journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *26*(4), 617–27. https://doi.org/10.1037/a0028666

Deci, E., & Ryan, R. (1995). A motivational approach to self: Integration in personality. In *Nebraska Symposium on Motivation: Vol. 38*. Lincoln, NE: University of Nebraska PRess.

Derogatis, L. R. (2000). *Brief Symptom Inventory (BSI) 18: Administration, Scoring, and Procedures Manual.* Minneapolis, MN: National Computer Systems.

Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, *90*(4), 286–297. https://doi.org/10.1037//0021-843X.90.4.286

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, *97*(2), 143–156. https://doi.org/10.1007/s11205-009-9493-y

Dobson, K. S., & Breiter, H. J. (1983). Cognitive assessment of depression: reliability and validity of three measures. *Journal of Abnormal Psychology*, *92*(1), 107–109. https://doi.org/10.1037/0021-843X.92.1.107

Edwards, V. J., Holden, G. W., Felitti, V. J., & Anda, R. F. (2003). Relationship between multiple forms of childhood maltreatment and adult mental health in community respondents: Results from the adverse childhood experiences study. *American Journal of Psychiatry*, *160*, 1453–1460.

Ernst, J. M., & Cacioppo, J. T. (1999). Lonely hearts: Psychological perspectives on loneliness. *Applied & Preventative Psychology*, *8*, 1–22.

Fals-Stewart, W., O'Farrell, T. J., & Hooley, J. M. (2001). Relapse among married or cohabiting substance-abusing patients: The role of perceived criticism. *Behavior Therapy*, *32*(4), 787–801. https://doi.org/10.1016/S0005-7894(01)80021-6

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347.

Feeney, B. C., & Collins, N. L. (2014). *A New Look at Social Support: A Theoretical Perspective on Thriving Through Relationships*. *Personality and Social Psychology Review*. https://doi.org/10.1177/1088868314544222

Felix, S. A. M., Hooley, J. M., & Hooker, C. I. (2017). Global perceived criticism ratings incrementally predict daily social appraisals over daily criticism exposure. *Manuscript Submitted for Publication*.

Felix, S. A. M., Masland, S. R., & Hooley, J. M. (2017). Is perceived criticism more than a relationship variable? *Manuscript Submitted for Publication*.

Fincham, F. D., & Rogge, R. (2010). Understanding relationship quality: Theoretical challenges and new tools for assessment. *Journal of Family Theory & Review*, *2*(4), 227–242. https://doi.org/10.1111/j.1756-2589.2010.00059.x

Fraley, R. C., Heffernan, M. E., Vicary, A. M., & Brumbaugh, C. C. (2011). The Experiences in Close Relationships-Relationship Structures questionnaire: a method for assessing attachment orientations across relationships. *Psychological Assessment*, *23*(3), 615–625. https://doi.org/10.1037/a0022898

Fraley, R. C., Hudson, N. W., Heffernan, M. E., Segal, N., Fraley, R. C., Hudson, N. W., … Segal, N. (2015). Are Adult Attachment Styles Categorical or Dimensional ? A Taxometric Analysis of General and Relationship-Specific Attachment Orientations. *Journal of Personality and Social Psychology*, *109*(2), 354–368.

Fraley, R. C., Waller, N. G., & Brennan, K. a. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, *78*(2), 350–365. https://doi.org/10.1037//0022-3514.78.2.350

Fritzsche, A., Dahme, B., Gotlib, I. H., Joormann, J., Magnussen, H., Watz, H., & Nutzinger, D. O. (2010). Specificity of cognitive biases in patients with current depression and remitted depression and in patients with asthma. *Psychological Medicine, 40*, 815–826. https://doi.org/10.1017/S0033291709990948

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology, 21*(4), 572–583. https://doi.org/10.1037/0893-3200.21.4.572

Gerlsma, C., van Duijn, M. a J., Hale, W. W., & van Hout, W. J. P. J. (2009). Perceived criticism: Associations with perceiver characteristics and interpersonal behaviour. *Psychiatry Research, 170*(2–3), 234–240. https://doi.org/10.1016/j.psychres.2008.09.007

Gleason, M. E. J., Iida, M., Shrout, P. E., & Bolger, N. (2008). Receiving support as a mixed blessing: Evidence for dual effects of support on psychological outcomes. *Journal of Personality and Social Psychology, 94*(5), 824–38. https://doi.org/10.1037/0022-3514.94.5.824

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology, 6*, 285–312. https://doi.org/10.1146/annurev.clinpsy.121208.131305

Gotlib, I. H., Kasch, K. L., Traill, S., Arnow, B. A., & Johnson, S. L. (2004). Coherence and specificity of information-processing biases in depression and social phobia. *Journal of Abnormal Psychology, 113*(3), 386–398. https://doi.org/10.1037/0021-843X.113.3.386

Gotlib, I. H., Krasnoperova, E., Neubauer Yue, D., & Joormann, J. (2004). Attentional biases for negative interpersonal stimuli in clinical depression. *Journal of Abnormal Psychology, 113*(1), 127–135. https://doi.org/10.1037/0021-843X.113.1.127

Hahlweg, K., & Conrad, M. (1983). Kategoriensystem zur Beobachtung Partnerschaftlicher Interaktion. Braunschweig, Germany: K. Hahlweg, Institut fu¨ r Psychologie, T. U. Braunschweig, Speilmannstrasse 12a, Braunschweig 38106, FRG.

Heim, C., & Binder, E. B. (2012). Current research trends in early life stress and depression: Review of human studies on sensitive periods, gene-environment interactions, and epigenetics. *Experimental Neurology, 233*(1), 102–11. https://doi.org/10.1016/j.expneurol.2011.10.032

Hendrick, S. S. (1988). A generic measure of relationship satisfaction. *Journal of Marriage and the Family*, *50*(1), 93–98. Retrieved from http://www.jstor.org/stable/352430

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. https://doi.org/10.2307/4615733

Hooker, C. I., Gyurak, A., Verosky, S. C., Miyakawa, A., & Ayduk, O. (2010). Neural activity to a partner's facial expression predicts self-regulation after conflict. *Biological Psychiatry*, *67*(5), 406–13. https://doi.org/10.1016/j.biopsych.2009.10.014

Hooley, J. M. (2007). Expressed emotion and relapse of psychopathology. *Annual Review of Clinical Psychology*, *3*, 329–52. https://doi.org/10.1146/annurev.clinpsy.2.022305.095236

Hooley, J. M., Gruber, S. A., Parker, H. A., Guillaumot, J., Rogowska, J., & Yurgelun-Todd, D. A. (2009). Cortico-limbic response to personally challenging emotional stimuli after complete recovery from depression. *Psychiatry Research: Neuroimaging*, *172*(1), 83–91. https://doi.org/10.1016/j.pscychresns.2009.02.001

Hooley, J. M., Gruber, S. A., Scott, L. A., Hiller, J. B., & Yurgelun-Todd, D. A. (2005). Activation in dorsolateral prefrontal cortex in response to maternal criticism and praise in recovered depressed and healthy control participants. *Biological Psychiatry*, *57*(7), 809–812. https://doi.org/10.1016/j.biopsych.2005.01.012

Hooley, J. M., & Miklowitz, D. J. (2017). Perceived criticism in the treatment of a high-risk adolescent. *Journal of Clinical Psychology*, *73*(5), 570–578. https://doi.org/10.1002/jclp.22454

Hooley, J. M., & Richters, J. E. (1991). Alternative measures of expressed emotion: a methodological and cautionary note. *Journal of Abnormal Psychology*, *100*(1), 94–7. https://doi.org/10.1037/0021-843X.100.1.94

Hooley, J. M., Siegle, G., & Gruber, S. A. (2012). Affective and Neural Reactivity to Criticism in Individuals High and Low on Perceived Criticism. *PLoS ONE*, *7*(9). https://doi.org/10.1371/journal.pone.0044412

Hooley, J. M., & Teasdale, J. D. (1989). Predictors of relapse in unipolar depressives: expressed emotion, marital distress, and perceived criticism. *Journal of Abnormal Psychology*, *98*(3), 229–235. https://doi.org/10.1037/0021-843X.98.3.229

Ibarra-Rovillard, M. S., & Kuiper, N. a. (2011). Social support and social negativity findings in depression: perceived responsiveness to basic psychological needs. *Clinical Psychology Review*, *31*(3), 342–52. https://doi.org/10.1016/j.cpr.2011.01.005

Jacobs, R. H., Reinecke, M. A., Gollan, J. K., & Kane, P. (2008). Empirical evidence of cognitive vulnerability for depression among children and adolescents: A cognitive science and developmental perspective. *Clinical Psychology Review*, *28*(5), 759–782. https://doi.org/10.1016/j.cpr.2007.10.006

Joiner, T. E., Metalsky, G. I., Lew, A., & Klocek, J. (1999). Testing the causal mediation component of Beck's theory of depression: Evidence for specific mediaiton. *Cognitive Therapy and Research*, *23*(4), 271–283. https://doi.org/10.1023/A:1018791531158

Joormann, J., & Gotlib, I. H. (2007). Selective attention to emotional faces following recovery from depression. *Journal of Abnormal Psychology*, *116*(1), 80–85. https://doi.org/10.1037/0021-843X.116.1.80

Joormann, J., Talbot, L., & Gotlib, I. H. (2007). Biased processing of emotional information in girls at risk for depression. *Journal of Abnormal Psychology*, *116*(1), 135–143. https://doi.org/10.1037/0021-843X.116.1.135

Joormann, J., Yoon, K. L., & Zetsche, U. (2007). Cognitive inhibition in depression. *Applied and Preventive Psychology*, *12*(3), 128–139. https://doi.org/10.1016/j.appsy.2007.09.002

Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: his and hers. *Psychological Bulletin*, *127*(4), 472–503. https://doi.org/10.1037/0033-2909.127.4.472

Kircanski, K., Mazur, H., & Gotlib, I. H. (2013). Behavioral activation system moderates self-referent processing following recovery from depression. *Psychological Medicine*, *43*, 1909–19. https://doi.org/10.1017/S0033291712002851

Kwon, J. H., Lee, Y., Lee, M. S., & Bifulco, A. (2006). Perceived criticism, marital interaction and relapse in unipolar depression - Findings from a Korean sample. *Clinical Psychology and Psychotherapy*, *13*(5), 306–312. https://doi.org/10.1002/cpp.500

Lakey, B., & Cassady, P. B. (1990). Cognitive processes in perceived social support. *Journal of Personality and Social Psychology*, *59*(2), 337–343. https://doi.org/10.1037//0022-3514.59.2.337

Lee, G., Barrowclough, C., & Lobban, F. (2014). Positive affect in the family environment protects against relapse in first-episode psychosis. *Social Psychiatry and Psychiatric Epidemiology*, *49*, 367–376. https://doi.org/10.1007/s00127-013-0768-x

Leff, J., & Vaughn, C. (1985). *Expressed Emotion in Families*. New York: Guilford.

Magaña, A. B., Goldstein, M. J., Karno, M., Miklowitz, D. J., Jenkins, J., & Falloon, I. R. H. (1986). A brief method for assessing expressed emotion in relatives of psychiatric patients. *Psychiatry Research*, *17*(3), 203–212. https://doi.org/10.1016/0165-1781(86)90049-1

Masland, S. R., Drabu, S., & Hooley, J. M. (n.d.). Is Percieved Criticism and independent construct? Evidence for discriminant validity across two samples. *Journal of Family Psychology*.

Masland, S. R., & Hooley, J. M. (2015). Perceived criticism: A research update for clinical practitioners. *Clinical Psychology: Science and Practice*, *22*(3), 211–222. https://doi.org/10.1111/cpsp.12110

Masland, S. R., Hooley, J. M., Tully, L. M., Dearing, K., & Gotlib, I. H. (2015). Cognitive-processing biases in individuals high on perceived criticism. *Clinical Psychological Science*, *3*(1), 3–14. https://doi.org/10.1177/2167702614529935

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology*, *1*, 167–95. https://doi.org/10.1146/annurev.clinpsy.1.102803.143916

Morrison, S. E., & Salzman, C. D. (2010). Re-valuing the amygdala. *Current Opinion in Neurobiology*, *20*(2), 221–230. https://doi.org/10.1016/j.conb.2010.02.007

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–9. https://doi.org/10.1016/j.tics.2005.03.010

Oliver, J. M., & Baumgart, E. P. (1985). The Dysfunctional Attitude Scale: Psychometric properties and relation to depression in an unselected adult population. *Cognitive Therapy and Research*, *9*(2), 161–167. https://doi.org/10.1007/BF01204847

Parker, G., Tupling, H., & Brown, L. B. (1979). A parental bonding instrument. *British Journal of Medical Psychology*, *52*(1), 1–10. https://doi.org/10.1111/j.2044-8341.1979.tb02487.x

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment*, *5*(2), 164–172. Retrieved from http://psycnet.apa.org/journals/pas/5/2/164/

Peterson-Post, K. M., Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2014). Perceived criticism and marital adjustment predict depressive symptoms in a community sample. *Behavior Therapy*, *45*(4), 564–575. https://doi.org/10.1016/j.beth.2014.03.002

Peterson, K. M., & Smith, D. A. (2010). To what does perceived criticism refer? Constructive, destructive, and general criticism. *Journal of Family Psychology*, *24*(1), 97–100. https://doi.org/10.1037/a0017950

Peterson, K. M., & Smith, D. A. (2011). Attributions for spousal behavior in relation to criticism and perceived criticism. *Behavior Therapy*, *42*(4), 655–666. https://doi.org/10.1016/j.beth.2011.02.005

Peterson, K. M., Smith, D. A., & Windle, C. R. (2009). Explication of interspousal criticality bias. *Behaviour Research and Therapy*, *47*(6), 478–486. https://doi.org/10.1016/j.brat.2009.02.012

R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Ausria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Rafaeli, E., Cranford, J. a, Green, A. S., Shrout, P. E., & Bolger, N. (2008). The good and bad of relationships: how social hindrance and social support affect relationship feelings in daily life. *Personality & Social Psychology Bulletin*, *34*(12), 1703–18. https://doi.org/10.1177/0146167208323742

Reis, H. T., Clark, M. S., & Holmes, J. G. (2004). Perceived partner responsiveness as an organizing construct in the sutdy of intimacy and closeness. In D. Mashek & A. P. Aaron (Eds.), *Handbook of closeness and intimacy* (pp. 201–225). Mahwah, NJ: Erlbaum. Retrieved from http://books.google.com/books?hl=en&lr=&id=viZ5AgAAQBAJ&oi=fnd&pg=PA201&dq=Perceived+partner+responsiveness+as+an+organizing+construct+in+the+sutdy+of+intimacy+and+closeness&ots=W3EfMSb2Vb&sig=2zPBgNtCt0EEbVewHzZYErbdhs0

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, *26*(4), 419–435.

Renshaw, K. D. (2007). Perceived criticism only matters when it comes from those you live with. *Journal of Clinical Psychology*, *63*(12), 1171–1179. https://doi.org/10.1002/jclp

Renshaw, K. D. (2008). The predictive, convergent, and discriminant validity of perceived criticism: A review. *Clinical Psychology Review*, *28*(3), 521–534. https://doi.org/10.1016/j.cpr.2007.09.002

Renshaw, K. D., Chambless, D. L., & Steketee, G. (2003). Perceived criticism predicts severity of anxiety symptoms after behavioral treatment in patients with obsessive-compulsive disorder and panic disorder with agoraphobia. *Journal of Clinical Psychology*, *59*(4), 411–421. https://doi.org/10.1002/jclp.10048

Richter, P., Werner, J., Heerlein, A., Kraus, A., & Sauer, H. (1998). On the Validity of the Beck Depression Inventory. *Psychopathology*, *31*(3), 160–168. https://doi.org/10.1159/000066239

Riso, L. P., Klein, D. N., Anderson, R. L., Ouimette, P. C., & Lizardi, H. (1996). Covergent and discriminant validity of perceived criticism from spouses and family members. *Behavior Therapy*, *27*(2), 129–137. https://doi.org/10.1016/S0005-7894(96)80010-4

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11392867

Scott, J., Colom, F., Pope, M., Reinares, M., & Vieta, E. (2012). The prognostic role of perceived criticism, medication adherence and family knowledge in bipolar disorders. *Journal of Affective Disorders*, *142*(1–3), 72–76. https://doi.org/10.1016/j.jad.2012.04.005

Seidman, G. (2012). Positive and negative: Partner derogation and enhancement differentially related to relationship satisfaction. *Personal Relationships*, *19*(1), 51–71. https://doi.org/10.1111/j.1475-6811.2010.01337.x

Sibley, C. G., Fischer, R., & Liu, J. H. (2005). Reliability and validity of the revised experiences in close relationships (ECR-R) self-report measure of adult romantic attachment. *Personality & Social Psychology Bulletin*, *31*(11), 1524–36. https://doi.org/10.1177/0146167205276865

Smith, D. A., & Peterson, K. M. (2008). Overperception of spousal criticism in dysphoria and marital discord. *Behavior Therapy*, *39*(3), 300–12.

https://doi.org/10.1016/j.beth.2007.09.002

Tarrier, N., Sommerfield, C., & Pilgrim, H. (1999). Relatives' expressed emotion (EE) and PTSD treatment outcome. *Psychological Medicine*, *29*(4), 801–11. https://doi.org/10.1017/S0033291799008569

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S.* (4th Editio). New York, NY: Springer.

von Polier, G. G., Meng, H., Lambert, M., Strauss, M., Zarotti, G., Karle, M., … Schimmelmann, B. G. (2014). Patterns and correlates of expressed emotion, perceived criticism, and rearing style in first admitted early-onset schizophrenia spectrum disorders. *The Journal of Nervous and Mental Disease*, *202*(11), 783–787. https://doi.org/10.1097/NMD.0000000000000209

Wearden, A. J., Tarrier, N., Barrowclough, C., Zastowny, T. R., & Rahill, A. A. (2000). A review of expressed emotion research in health care. *Clinical Psychology Review*, *20*(5), 633–666. https://doi.org/10.1016/S0272-7358(99)00008-2

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale— Revised*. New York: Psychological Corporation.

Weissman, A. N. (1980). *Assessing depressogenic attitudes: A validation study.* Paper presented at the 51st Annual Meeting of the Eastern Psychological Association, Hartford.

Weissman, A. N., & Beck, A. T. (1978). Development and validation of the dysfunctional attitudes scale: Preliminary investigation. In *Annual Meeting of the American Educational Research Association*. Chicago, November. Retrieved from http://files.eric.ed.gov/fulltext/ED167619.pdf

Wells, T. T., Vanderlind, W. M., Selby, E. A., & Beevers, C. G. (2014). Childhood abuse and vulnerability to depression: cognitive scars in otherwise healthy young adults. *Cognition & Emotion*, *28*(5), 821–833. https://doi.org/10.1080/02699931.2013.864258

White, J. D., Strong, J. E., & Chambless, D. L. (1998). Validity of the Perceived Criticism Measure in an undergraduate sample. *Psychological Reports*, *83*(1), 83–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9775664

Wilhelm, K., Niven, H., Parker, G., & Hadzi-Pavlovic, D. (2005). The stability of the Parental Bonding Instrument over a 20-year period. *Psychological Medicine*, *35*(3), 387–393. https://doi.org/10.1017/S0033291704003538

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and emotional disorders*. Chichester: John Wiley and Sons Ltd.