INTERSPOUSAL CRITICALITY BIAS IN DEPRESSION AND MARITAL DISCORD

A Thesis

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Master of Arts

by

Kristina M. Peterson, B.A.

David A. Smith, Director

Graduate Program in Psychology
Notre Dame, Indiana
July 2007
© Copyright by

KRISTINA M. PETERSON

2007

All rights reserved
INTERSPOUSAL CRITICALITY BIAS IN DEPRESSION AND MARITAL DISCORD

Abstract

By

Kristina M. Peterson

Reports of perceived spousal criticism better predict relapse following recovery from major depressive disorder than do measurements of “actual” spousal criticism from Expressed Emotion interviews coded by outside raters (Hooley & Teasdale, 1989). This study examines relations between perceived and “actual” criticism in depressed and non-depressed spouses. Married couples were evaluated using structured interviews, questionnaires, and an observationally-coded dyadic interaction. Signal detection methods assessed sensitivity and bias in reports of perceived criticism. Results suggested that depressed and maritaly discordant spouses are biased towards perceiving spousal criticism, suggesting that criticality bias reflects a latent cognitive style underlying depression and marital discord. That bias was related to reports of global perceived criticism suggests that spousal reports of perceived criticism capture both actual criticism and reporters’ cognitive bias towards over- or under-perceiving criticism. Reports of
perceived criticism may better predict relapse from depression than ratings of actual criticism in part because they capture such cognitive tendencies.
TABLES ........................................................................................................... iii

ACKNOWLEDGEMENTS .............................................................................. iv

1. INTRODUCTION ......................................................................................... 1
1.1 Depression, Criticism, and Marital Adjustment ........................................... 1
1.2 Reports of Criticism .................................................................................. 4
1.3 Types of Criticism: Valence ....................................................................... 7
1.4 Signal Detection Theory ........................................................................... 8
1.5 The Present Investigation .......................................................................... 9

2. METHOD ...................................................................................................... 12
2.1 Participants ............................................................................................... 12
2.2 Procedure ................................................................................................. 14
2.3 Observational Measures .......................................................................... 15
2.4 Self-Report Measures .............................................................................. 18

3. RESULTS ..................................................................................................... 20
3.1 Part One: Global Criticism ....................................................................... 21
3.2 Part Two: Laboratory Criticism ................................................................. 23
3.3 Part Three: Signal Detection Analyses of Discussion Data ......................... 27
3.4 Part Four: Relationships between Criticism Assessments ......................... 33

4. DISCUSSION ............................................................................................... 36
4.1 Depression, Marital Adjustment, and Criticality Bias ................................. 36
4.2 Global Perceived Criticism ....................................................................... 41
4.3 Types of Criticism .................................................................................... 44
4.4 Limitations and Future Directions ............................................................. 45

APPENDICIES ............................................................................................... 48

REFERENCES ................................................................................................. 52
TABLES

1. MEANS AND STANDARD DEVIATIONS OF MARITAL ADJUSTMENT AND CRITICISM ............................................................ 22

2. MEANS (STANDARD DEVIATIONS) OF LABORATORY CRITICISM BROKEN DOWN BY INTENSITY LEVEL.......................... 25

3. MEANS (STANDARD DEVIATIONS) OF TYPE OF CRITICISM AND OVERALL VALENCE DURING THE LABORATORY DISCUSSION ................................................................. 27

4. MEANS (STANDARD DEVIATIONS) OF SIGNAL DETECTION INDICES FOR OVERALL CRITICISM.............................................. 29

5. ZERO-ORDER CORRELATIONS BETWEEN PCRITGLOBAL AND OTHER CRITICISM MEASURES WITHIN GROUPS .................... 34

6. ZERO-ORDER CORRELATIONS BETWEEN PCRITGLOBAL AND OTHER PERCEIVED CRITICISM MEASURES ACROSS GROUPS ................................................................. 35
ACKNOWLEDGMENTS

I thank Dr. David Smith, my faculty advisor, for assisting me on this project and for his support throughout this process. I also thank Dr. Thomas Merluzzi, Dr. Scott Monroe, and Dr. Gerald Haeffel for their helpful comments and for serving on my committee. Finally, I thank the students in the Marital Research Laboratory for their help in data collection, coding, and entry.

This research was funded by NIMH RO3 MH066301-02.
CHAPTER 1

INTRODUCTION

People who are feeling depressed often claim that their spouses are highly critical of them (Hooley & Teasdale, 1989). It is unclear, however, how much these reports reflect the generally negative views taken by people who are feeling depressed (Ingram, Miranda, & Segal, 1998) or exposure to genuinely hypercritical spouses. The main purpose of this study is to examine the possibility that those who are depressed over-report the actual criticism they receive from their spouse. This will be accomplished using a design that permits objective assessment of exposure to interspousal criticism against which reports of perceived criticism can be compared. A secondary purpose of this study is to clarify the meaning of criticism by breaking it down into different types based on both valence and intensity. We will examine the type of criticism depressed spouses are subject to in comparison to non-depressed control spouses and will examine the correlation between these types of criticism and global perceptions of criticism.

1.1 Depression, Criticism, and Marital Adjustment

Criticism was first linked to depression in studies of relapse and Expressed Emotion (EE: Vaughn & Leff, 1976a; Hooley, Orley, & Teasdale, 1986; Hooley & Teasdale, 1989). EE, which is indicated by family members’ critical, hostile, or
emotionally over-involved statements about their relatives, has traditionally been assessed with the Camberwell Family Interview (Brown & Rutter, 1966; Rutter & Brown, 1966; Vaughn & Leff, 1976b) or with the Five Minute Speech Sample (Magaña, et al., 1986) in which relatives’ critical comments stated in the absence of the patient are coded by outside observers.

The importance of EE is underscored by a study of potential demographic and clinical relapse predictors in a sample of patients hospitalized for unipolar depression (Hooley et al., 1986). Among the candidate demographic and clinical variables tested, only EE predicted relapse to a statistically significant degree. A meta-analysis of this literature suggests that patients who are hospitalized for the treatment of depression have a 2 to 3 times greater risk of relapse when they live with “high EE” relatives than when they live with “low EE” relatives (Butzlaff & Hooley, 1998).

Although EE predicts poor outcomes in depression and a number of other psychological and physical conditions, such as schizophrenia, eating disorders, rheumatoid arthritis, and diabetes (see Wearden, Tarrier, Barrowclough, Zastowny, & Rahill, 2000 for a review), the link between EE and relapse may be especially strong in depression. A meta-analysis of studies examining EE and several psychological disorders revealed that EE was a significantly better predictor of relapse in mood disorders than in schizophrenia (Butzlaff & Hooley, 1998). Likewise, relapse in depression is associated with lower levels of criticism than is relapse in schizophrenia, suggesting that depressed people may be particularly sensitive to criticism (Hooley et al., 1986; Vaughn & Leff, 1976a).
Additionally, *perceived* criticism may be especially strong in predicting relapse. When EE, marital satisfaction, and responses to the question, “How critical is your spouse of you?” were compared, the amount of criticism depressed patients perceived accounted for more of the variance in relapse than both marital satisfaction and EE combined (Hooley & Teasdale, 1989). Therefore, perceived criticism, rather than observer ratings of EE, may be especially strong predictors of relapse, possibly because it reflects how much criticism is “getting through” to patients (p. 234; Hooley & Teasdale, 1989).

Criticism is a generic stressor which has been linked not only to depression but also to marital dysfunction (Hayhurst, Cooper, Paykel, Vernals, & Ramana, 1997; Hooley & Teasdale, 1989; Lynch, Robins, & Morse, 2001). As part of the definition of the “demandingness” dimension of the demand-withdraw pattern, criticism is related to poor marital adjustment (Lynch, Robins, & Morse, 2001) and distinguishes between distressed and non-distressed couples (Christensen & Shenk, 1991). Gottman names criticism as one of the “Four horsemen of the Apocalypse”, factors he believes to be especially predictive of divorce (Gottman, 1994). Furthermore, according to the marital discord model of depression, poor marital adjustment may lead to depression in part because of reduced spousal support and increased stress and criticism (see Beach, Sandeen, & O’Leary, 1990). Therefore, there is both strong empirical and theoretical evidence for the relationship between poor marital adjustment and spousal criticism.

Not only are both depression and marital adjustment each related to spousal criticism, depression and marital adjustment are also related to each other (see Beach, Smith, & Fincham, 1994; Gotlib & Hammen, 1992; Whisman, 2001 for reviews). A
recent meta-analysis revealed that approximately 44 percent of the variance in depression can be explained by level of marital adjustment (Whisman, 2001). Longitudinally, spouses who are dissatisfied in their marriages at baseline are approximately three times as likely to experience a major depressive episode in the following year (Whisman & Bruce, 1999). Likewise, first episodes of major depressive disorder are often preceded by marital and other interpersonal stressors (Whisman & Bruce, 1999), and marital dissatisfaction is a predictor of relapse in those who are hospitalized for depression (Hooley & Teasdale, 1989). Processes such as interspousal criticism that relate to both depression and marital discord are especially interesting because they might account for at least part of the substantial relationship between depression and marital adjustment (Dehle & Weiss, 1998; O’Leary, Christian, & Mendell, 1994; Scott & Cordova, 2002; see Beach et al., 1994; Gotlib & Hammen, 1992; Whisman, 2001 for reviews).

1.2 Reports of Criticism

Although the picture emerging from this brief summary of the criticism literature is fairly consistent in that criticism is related to both marital discord and depression, it is worth remaining mindful of a potentially important but underemphasized distinction between criticism as it is perceived and criticism as it is expressed. In the marital literature, observational coding studies are concerned with expressed criticism. However, demand-withdraw and other marital assessments of criticism are self-report and therefore concerned with perceived criticism (Christensen & Shenk, 1991; Fiscella & Campbell, 1999; Lynch et al., 2001). In the depression literature, EE ratings, on the one hand, are made from observations of criticism (e.g., expressed criticism) during an interview with a
family member. Perceived criticism, on the other hand, is a perceptual phenomenon assessed via self-report. Combining these disparate literatures and assessment techniques requires endorsement of the untested assumption that perceived and expressed criticism, as it is observed by either outside raters or reported by spouses, are related.

Although there are no known studies to date directly comparing the agreement between perceived and expressed criticism, there are a few studies comparing observer ratings of criticism from EE interviews conducted in the absence of the patient and patient reports of the amount of criticism they perceive from their spouse in general. In one study, perceived criticism was not significantly related to spouses’ own reports of criticism, but it was modestly correlated with observer ratings of EE, thereby lending some support to the notion that depressed patients do, in fact, accurately perceive criticism from spouses (Hooley & Teasdale, 1989). In another study, however, even though perceived spousal criticism and marital discord were related to each other, expressed criticism as measured in an EE interview was not related to either one (Hayhurst et al., 1997). In sum, the few studies that have examined the relationship between perceived and expressed criticism have yielded mixed results and therefore distinctions between different reporters of criticism and different criticism assessment techniques deserve further clarification.

Especially in the case of depression, where a general negativity bias might influence perceptions, it is probably unwise to assume that perceived and expressed criticism are directly related. It is likely that the general negative affectivity associated with depression is related to a negative outlook and possibly even a specific oversensitivity to criticism (Clark & Watson, 1991; Watson & Clark, 1984). Indeed,
studies have shown that depressed people interpret ambiguous situations negatively (Lawson, MacLeod, & Hammond, 2002; Nunn, Matthews, & Trower, 1997); attend more to negative interpersonal cues, such as facial expressions of sadness (Gotlib, Krasnoperova, Yue, & Joormann, 2004); and are biased away from identifying happy facial expressions (Surguladze et al., 2004). Furthermore, depressed people may be more likely to perceive negative environmental stimuli than non-depressed people. For example, Roth and Rehm (1980) found that while watching a videotaped interaction of themselves in a role-play with an experimenter, depressed inpatients identified significantly more of their own negative behaviors and significantly less of their own positive behaviors than non-depressed inpatient controls even when actual levels of negative and positive behaviors were controlled for via outside observer ratings. It is important to note however, that significant differences were observed between subjects’ ratings and outside observer ratings; in fact, the only non-significant difference between ratings was the comparison between outside observers’ ratings and depressed subjects’ ratings of their own negative behaviors (Roth & Rehm, 1980). Thus, these results suggest that both depressed and non-depressed inpatients were inaccurate in their ratings of their behaviors and that depressed patients tended to have a bias towards identifying more of their negative behaviors and less of their positive behaviors than non-depressed patients. Therefore, depressed people may have a bias towards identifying negative stimuli and thus they may be more likely to perceive and recall negative aspects of their relationships, such as instances of spousal criticism (Gotlib & Krasnoperova, 1998).
1.3 Types of Criticism: Valence

Since depressed people appear particularly attuned to negative stimuli, it may be important to take into account an underappreciated aspect of criticism, that is, the valence of criticism. In the EE literature, only criticism that is delivered in a negative tone is coded as criticism (Vaughn & Leff, 1976b). However, criticism delivered in a positive tone may be important as well. Most people can recall times when criticism helped them and was appreciated, such as criticism about something they could improve. In this sense, constructive criticism may have positive effects and may even buffer the effects of destructive criticism. Furthermore, similar to Roth and Rehm’s (1980) finding of a differential response bias to positive and negative behaviors, depressed people may be biased towards perceiving negative criticism and towards under-perceiving positive criticism in comparison to non-depressed controls. Depressed people may also interpret most constructive criticism as destructive. Indeed, one study involving spousal perceptions and marital distress has indicated that although distressed and non-distressed couples did not differ in the intent of their messages in a marital interaction, spouses in discordant marriages were perceived as more negative by their partner (Gottman et al., 1976). Given the relationship between depression and marital distress, depressed people may also perceive their spouses as acting more negatively than was intended. In sum, especially when depression is involved, it may be useful to consider not only the frequency of criticism but also the valence associated with it.
1.4 Signal Detection Theory

While depressed people may be biased in the sense that they seem to perceive and recall more negative and less positive information than do non-depressed people, it is less clear whether depressed and non-depressed people can discriminate between negative and nonnegative information. That is, despite a fair amount of work on depressive bias, there is relatively little evidence regarding depressive “sensitivity.” In the proposed study, methods from signal detection theory (Green & Swets, 1966; Macmillan & Creelman, 2005) will be used to separate the degree of bias from the degree of sensitivity with which people perceive spousal criticism.

Signal detection methods can be used to compare subjects’ ratings of the presence or absence of a signal with whether the signal was actually present or not. In this study, the actual presence of a signal can be measured from two different perspectives: the “insider” perspective and the “outsider” perspective (e.g., Floyd & Markman, 1983). The “insider” perspective will compare perceived criticism with spouses’ ratings of intended criticism. The “outsider” perspective will compare perceived criticism with outside judges’ ratings of observed criticism. Observed criticism will be used as an additional perspective because spouses may be inaccurate or biased in reporting the criticism they intend. That is, they may not realize they are being critical when actually they are. Therefore, two different signal detection analyses will compare (a) perceived and intended and (b) perceived and observed ratings of criticism.

Comparisons of criticism reports give rise to the two main indices of signal detection, sensitivity and bias. Sensitivity reflects subjects’ ability to accurately distinguish between when a signal is present and when it is not; in other words, it reflects
the difference between the hit and false alarm rates. High sensitivity indicates high hit and low false alarm rates whereas low sensitivity indicates either both high hit and false alarm rates or low hit and false alarm rates. Bias, a measure independent of sensitivity, reflects the tendency to use “yes” or “no” responses; in other words, it reflects the sum of the hit and false alarm rates. While the terminology may seem counterintuitive, a negative bias would indicate both high hit and false alarm rates whereas a positive bias would indicate low hit and false alarm rates. In sum, the use of signal detection methods in this study will allow us to compare ratings of criticism to determine subjects’ degree of sensitivity and response bias in rating perceived criticism.

1.5 The Present Investigation

Part one: Global perceived criticism. The first purpose of this study is to examine global reports of criticism (e.g., Hooley & Teasdale, 1989) in relation to depression and marriage. Specifically, we hypothesize that: (1a) depressed participants, in comparison to non-depressed participants, will have a greater discrepancy between their own reports of perceived criticism and their spouse’s reports of intended criticism. Specifically, this heightened discrepancy will be produced by (1b) higher levels of perceived criticism in the depressed group. Finally, we hypothesize that (1c) depressed spouses will perceive significantly more criticism than non-depressed spouses, even after controlling for marital adjustment.

Part two: Laboratory criticism. After investigating overall levels of perceived criticism via self-report questionnaires, we will further investigate spousal criticism as rated by study participants after watching a replay of a videotaped discussion that they
had with their spouse. Specifically, we will test the hypothesis that, as with the global measure of criticism, in data based on ratings of themselves interacting with their spouses, (2a) there will be a significant interaction between group and reporter of criticism with (2b) depressed participants perceiving more criticism than control participants.

Since ratings of the videotaped discussion will include not only reports of criticism (on a 0-3 scale of differing intensity and duration) but also of valence (using a positive, neutral, negative rating), criticism in the laboratory task will also be explored in relation not only to severity but also in relation to valence. With regards to severity, we will explore specific levels of criticism (1-3) to see if depressed participants differ from controls on the frequency and intensity of criticism to which they are exposed. As far as valence is concerned, we will examine destructive and constructive criticism to test the hypotheses that depressed participants will perceive (2c) more destructive and (2d) less constructive criticism than control participants.

**Part three: Signal detection analyses of discussion data.** The third purpose of the study is to compare reports of criticism made by spouses to those made by outside observers (via signal detection methods) during a videotaped discussion. We hypothesize that depressed people are (3a) less sensitive (viz., less able to discriminate between critical and non-critical comments) and (3b) more biased in perceiving a criticism from their spouse (viz., they tend to report criticism rather than to not report it). If these hypotheses are supported, we will also test whether depressed spouses are (3c) less sensitive and (3d) more biased in perceiving spousal criticism even after controlling for
marital adjustment since marital adjustment may also be associated with sensitivity and bias.

Next, since criticism in EE studies is only coded if it is negative in nature, we will isolate instances of negative or destructive criticism to create additional sensitivity and bias scores. We hypothesize that: depressed spouses are (3e) less sensitive and (3f) more biased in their perceptions of destructive criticism relative to constructive criticism.

Part four: Relationships between criticism assessments. Finally, to clarify the relationship between reports of global perceived criticism and reports of specific instances of criticism during the marital discussion, exploratory correlational analyses will be conducted. Within groups, we will test the heretofore untested assumption that global perceived criticism is related to expressed criticism (as measured by spouses and outside observers). We hypothesize that (4a) in both groups global perceived criticism will be correlated with criticism during the laboratory discussion. Additionally, across groups we hypothesize that (4b) bias in perceiving criticism in the interaction task will be associated with global ratings of perceived criticism. We will further explore the global perceived criticism item via its correlations with the types of criticism (based on severity and valence) measured in the laboratory discussion. Since these analyses are exploratory, there are no specific hypotheses for these analyses.
CHAPTER 2

METHOD

2.1 Participants

We recruited married couples from South Bend, IN and surrounding areas. Control couples were recruited through newspaper advertisements for a “marital communication” study. Depressed couples were recruited through referrals from physicians, newspaper ads, and fliers placed in a local outpatient mental health clinic targeting people who are “feeling depressed.” Most couples were paid $50 for their participation. However, because the depressed participants took a substantially longer amount of time to complete the study, payment was increased to $100 for some of the depressed couples.

To formally categorize couples into groups on the basis of diagnostic status, we used the Structured Clinical Interview for DSM-IV-TR Axis I Disorders- Non-Patient Edition (SCID-I/NP, Feb 2001 revision: First, Gibbon, Spitzer, & Williams, 2001; Spitzer, Williams, Gibbon, & First, 1992; Williams et al., 1992). If either spouse met criteria for current major depressive disorder according to the Diagnostic and Statistical Manual of Mental Disorders (text revision; DSM-IV-TR; American Psychiatric Association, 2000) criteria, the couple was placed in the depressed group. Control couples were those in which neither spouse met criteria for current depression. The final
sample consisted of 108 couples (31 depressed and 77 control couples). Couples were excluded from analyses if either partner was acutely suicidal, had experienced psychotic symptoms, or if they met criteria for substance abuse or dependence within the past year, post-traumatic stress disorder within the past two years, obsessive-compulsive disorder within the last two years, or had current or past manic episodes. 59 couples from the original recruited sample of 168 were excluded on the basis of these criteria.

The majority of the demographic information for participants did not differ between control and depressed groups. The majority of participants gave their race/ethnicity as White (77.8%). 9.3% indicated they were African American, 3.7% indicated Hispanic, 0% indicated Asian, and 5.0% gave another race/ethnicity (e.g., “mixed”). Six participants did not indicate a race/ethnicity. Couples had been married for an average of 13.7 (SD = 13.83) years. Although, the length of marriages did not differ between groups, the depressed spouses had been married significantly more times (M = .60, SD = .89) than control participants (M = .25, SD = .57). Spouses averaged 42.29 years of age (SD = 13.82). Couples averaged 2.09 children (SD = 1.62; range = 0 – 7). Annual family incomes ranged from $0 to $150,000, with a median of $47,500 (M = 48,243, SD = 29,378). Although incomes were the same as reported by both depressed and control spouses, the number of years of education differed between groups. Depressed spouses had significantly less years of education (M = 12.80, SD = 1.89) than the control spouses (M = 14.68, SD = 2.97).
2.2 Procedure

This study was conducted under the oversight of the University of Notre Dame Institutional Review Board. All participants were treated in accordance with the ethical standards of the American Psychological Association. Couples participated in a 3 hour laboratory visit conducted by trained graduate and undergraduate psychology students. According to the study protocol, participants completed questionnaires, engaged in a 10-minute videotaped desired change interaction task (e.g., Pasch & Bradbury, 1998), and provided ratings of the videotaped discussion task. Participants were assessed with a structured diagnostic interview, as described previously, which was conducted by trained graduate students (SCID-I/NP, Feb 2001 revision: First et al., 2001; Spitzer et al., 1992; Williams et al., 1992). All interviews were videotaped. A second diagnostician reviewed 80 SCID interviews (approximately 23% of interviews), establishing a $K = .82$ level of diagnostic agreement.

For the desired change interaction task, spouses met together with an experimenter who informed couples who the “target” of the discussion would be. For control couples, the target spouse was randomly chosen via a coin toss. For depressed couples, the depressed spouse was designated as the target. For expository purposes, we will refer to the non-target spouse as the “source”, though participants were not aware of this designation.

Each spouse then independently listed five things they would like changed about the target spouse and ranked these topics in order of importance. At the point of generating topics, spouses understood that one of the topics would be chosen for a subsequent videotaped discussion. The experimenter then chose the highest ranked item
appearing on both lists as the topic for discussion. If there were no overlapping items, the experimenter asked permission to disclose the target’s list to the source spouse who then chose a topic from the target’s list. Spouses did not otherwise see or discuss each other’s lists of topics. Couples were then videotaped engaging in a 10 minute discussion of the desired target spouse change. Two identically time-stamped videotapes were made of this 10 minute interaction for subsequent independent rating by each spouse.

Following the discussion, spouses were taken to separate rooms where they each watched videotaped instructions explaining the rating task. After subjects indicated understanding of the task, a time-stamped copy of the videotaped interaction was replayed while subjects listened with headphones (so that the experimenters could not hear the couple’s conversation). During playback, experimenters paused each spouse’s video following the same 30-second intervals. While the videos were paused, target participants completed a questionnaire assessing their perceived criticism and perceived valence, whereas source participants completed a questionnaire assessing their intended criticism and intended valence. Finally, outside observers also rated couples videotaped interactions for the criticism and valence of the source spouse following the same 30-second intervals as participants. In this manner, three perspectives (e.g., perceived, intended, and observed) on the same interaction were obtained.

2.3 Observational Measures

Participants’ ratings of their interaction. After completing a videotaped discussion with their spouse, participants independently watched a time stamped copy of their interaction, and at each 30-second interval during the discussion, they completed a
measure assessing (a) criticism, (b) valence, (c) their thoughts/feelings, and (d) their spouse’s thoughts/feelings (see Appendices A and B). Specifically, target spouses responded to the questions, “How critical was your partner being of you?” and “Was your partner being positive, neutral, or negative towards you?”; thereby rating both perceived criticism and valence. Source spouses responded to the questions “How critical were you intending to be of your partner?” and “Were you being positive, negative, or neutral towards your partner?”; thereby rating intended criticism and valence.

Perceived and intended criticism were rated on the following scale: 0 = “not at all critical”, 1 = “slightly critical”, 2 = “moderately critical”, and 3 = “very critical.” Participants were instructed to consider both the duration and intensity of the criticism over the entire 30-second interval. As such, the ratings that both participants and judges used were based on similar criteria. Likewise, participants rated valence in a manner similar to the judges where 1 = positive, 2 = neutral, and 3 = negative valence.

*Ratings of types of criticism.* Ideally, participants would rate critical comments that their spouses made as either constructive, neutral, or destructive. However, in this study we were confined to valence ratings as indicators of constructive, neutral, or destructive criticism because spouses did not directly classify critical comments as constructive, destructive, or neutral but did so indirectly through the valence (e.g., positive, neutral, or negative) attached to comments during the 30-second intervals. Thus, for our purposes, constructive criticism was operationalized as those instances where there was criticism and the corresponding valence was positive. Likewise, destructive criticism was indicated by instances of criticism rated with a negative valence. This
approach provides ratings that take into account both the comment and the context in which it was stated.

**Judges’ ratings of observed valence and criticism.** While spousal ratings of intended criticism provide an “insider” perspective and can be used to tap into couples’ private style of communication that only they might understand, outside observer ratings may also be useful because spouses’ may be inaccurate in their ratings of intended criticism in that they may be critical without such intent. Therefore, three outside judges independently rated couples’ videotaped interactions at each of the same 30-second intervals for valence and critical comments made by the source participant.

Training of most coders began with reading primary source articles that provided definitions of criticism as it was instantiated via other major observational coding systems (e.g., Magaña et al, 1986; Vaughn & Leff, 1976b; Weiss & Summers, 1983). Specifically, coders were trained to consider as critical any comment that expressed dislike, disapproval, or resentment of the spouse’s personality or behavior. Criticism ratings were made following each 30-second interval using a “0” to “3” scale: where “0” = “no criticism”; “1” = “low intensity and short duration criticism”; “2” = “high intensity or long duration criticism”; and “3” = “high intensity and long duration criticism, or global criticism of character.” Valence was rated on the following 1 to 3 scale: 1 = positive, 2 = neutral, and 3 = negative tone towards their spouse. The 30-second intervals used by judges matched those used by spouses, and the criticism and valence rating scales were the same as the spouses’ scales. Prior to coding study tapes, sample tapes were coded in groups, with coders discussing their ratings out-loud in order to expose and
resolve misunderstandings and to clarify definitions. Coders then rated study tapes independently of the other coders.

In all, seven judges provided criticism ratings. For each set of 25 videos, three judges were randomly assigned to code for criticism and valence. The reliability of these judges on each of the 30 second intervals was calculated with an interclass correlation coefficient, $ICC(3, k)$, in order to correspond to the clip-by-clip level at which subsequent signal detection analyses were undertaken (Shrout & Fleiss, 1979). The average reliability of these ratings across the twenty clips for was .71 for the criticism ratings and .69 for the valence ratings.

2.4 Self-Report Measures

Marital adjustment. Marital adjustment was measured with the Dyadic Adjustment Scale (see Appendix C), a 32-item scale assessing marital satisfaction, cohesion, consensus, and affective expression (DAS; Spanier, 1976). Scores on the DAS range from 0 to 151, with lower scores indicating less marital adjustment (or more marital discord). The DAS is a widely used measure with excellent internal consistency reliability (> .90) as well as compelling evidence of convergent and discriminant validity (Heyman, Sayers, & Bellack, 1994; Kazac, Jarmas, & Snitzer, 1988; Spanier, 1976). The DAS is highly correlated with other measures of relationship conflict, such as the Quality of Marriage Index (Heyman et al., 1994). It also distinguishes married from divorced couples and scores change with improvements made during marital therapy (Spanier, 1976; Whisman & Jacobson, 1992).
The DAS mean and standard deviation for target spouses are presented in Table 1. Cronbach’s alpha estimate of the internal consistency of the DAS was .94 for the target spouses in this sample. Using the conventional cutoff of 98, which has been adopted because of evidence that it maximally separates distressed clinic couples from non-distressed community couples (Eddy, Heyman, & Weiss, 1991), 24.0% of target participants in the non-depressed group and 53.3% of target participants in the depressed group were maritally distressed.

*Spousal criticism.* A modified version of the Perceived Criticism Scale developed by Hooley and Teasdale (1989) assessed global criticism. The scale consists of two questions: (a) “How critical is your spouse of you?”, and (b) “How critical are you of your spouse?” The original scale ranged from 1 (*not at all critical*) to 10 (*very critical indeed*), but for the purpose of the current study the scale was changed to a 0 (*not at all critical*) to 5 (*extremely critical*) scale to correspond with a larger scale, the General Criticism Scale (see Appendix D; Couture, 2001; Couture, Smith, & Montemarano, 2000; Smith, Couture, & Myers, 2001) that incorporates multiple perspectives (e.g., self-criticism and criticism from others) assessed on a 0 to 5 scale.

The Perceived Criticism scale has demonstrated good temporal stability and concurrent validity with the Camberwell Family Interview (Van Humbeeck, Van Audenhove, De Hert, Pieters, & Storms, 2002). Likewise, it has evidenced discriminant validity by its lack of correlation with measures of depression or personality traits (Riso, Klein, Anderson, Ouimette, & Lizardi, 1996). The means and standard deviations of the Perceived Criticism scale are presented in Table 1.
CHAPTER 3

RESULTS

Results are presented in four parts describing: (1) global criticism, (2) laboratory criticism, (3) signal detection analyses of discussion data, and (4) the relationship between the types of criticism assessments developed in parts 1-3.

In part one, we present analyses concerning the self-report measure of global perceived and intended criticism (Hooley & Teasdale, 1989). The global perceived criticism item “How critical is your spouse of you?” is referred to as “PCrit_{global}” and the global intended criticism item “How critical are you of your spouse?” is referred to as “ICrit_{global}.” Because all spouses completed both the perceived and intended criticism assessments, in order to reduce dependency in the data that results from having married couples as participants (Kenny, 1995), PCrit_{global} refers only to target spouses’ reports whereas ICrit_{global} refers only to source spouses’ reports.

In part two, we present the micro-analytic indicators of criticism expressed during the laboratory discussion from three different perspectives. Here, “PCrit_{lab}” denotes target spouse reports of their perceptions of criticism received from their source partners during the laboratory task. “ICrit_{lab}” denotes source spouse reports of their intended criticism towards their target partners during the laboratory task. “OCrit_{lab}”
denotes outside observer ratings of criticism as expressed by the source spouse toward the target spouse during the laboratory task.

3.1 Part One: Global Criticism

The means and standard deviations of the main criticism measures, including \( PC_{\text{global}} \) and \( IC_{\text{global}} \), are presented in Table 1 for the control and depressed groups. Inspection of the means and standard deviations reveals that in both groups, targets perceived their spouse as globally only slightly critical; source spouses were rating their own intended criticism at similar levels, suggesting a correspondence between what spouses perceive and intend.

To test hypothesis 1a, that there would be a larger discrepancy between reports of perceived and intended criticism in the depressed group in comparison to the non-depressed group, we subjected the global criticism items to a 2 (Group: control vs. depressed) \( \times 2 \) (Perspective: perceived vs. intended) split plot ANOVA, with Group as a between subjects factor and Perspective as a within subjects factor. Contrary to hypothesis 1a, there was not a significant Group \( \times \) Perspective interaction, \( F(1, 102) = 1.30, ns \). Furthermore, there was not a significant main effect of Group, \( F(1, 102) = .87, ns \) nor a main effect of Perspective, \( F(1, 102) = 2.42, ns \). Therefore, it appears that depressed spouses perceived and were subjected to levels of criticism similar to their non-depressed counterparts. What is more, when spouse reports of perceived and intended criticism within couples were compared, reports within depressed couples did not differ more than reports within control couples suggesting that depressed spouses were not over- or under-estimating criticism more than those in the non-depressed group.
TABLE 1
MEANS AND STANDARD DEVIATIONS OF MARITAL ADJUSTMENT AND CRITICISM

<table>
<thead>
<tr>
<th>Measure</th>
<th>Control</th>
<th>Depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Marital adjustment</td>
<td>109.05</td>
<td>15.72</td>
</tr>
<tr>
<td>PCrit&lt;sub&gt;global&lt;/sub&gt;</td>
<td>2.00</td>
<td>1.45</td>
</tr>
<tr>
<td>ICrit&lt;sub&gt;global&lt;/sub&gt;</td>
<td>2.43</td>
<td>1.17</td>
</tr>
<tr>
<td>PCrit&lt;sub&gt;lab&lt;/sub&gt;</td>
<td>.62</td>
<td>.71</td>
</tr>
<tr>
<td>ICrit&lt;sub&gt;lab&lt;/sub&gt;</td>
<td>.73</td>
<td>.72</td>
</tr>
<tr>
<td>OCrit&lt;sub&gt;lab&lt;/sub&gt;</td>
<td>.14</td>
<td>.19</td>
</tr>
</tbody>
</table>

NOTE: PCrit<sub>global</sub> and ICrit<sub>global</sub> refer to reports of perceived and intended global criticism on the self-report questionnaire. PCrit<sub>lab</sub>, ICrit<sub>lab</sub>, and OCrit<sub>lab</sub> refer to perceived, intended, and observed criticism from the laboratory discussion. Except for ICrit<sub>global</sub> and ICrit<sub>lab</sub>, which are reported by source spouses and OCrit<sub>lab</sub>, which is reported by outside raters, all other variables are reported by target spouses.

In addition to the non-significant effects above, a one-way ANOVA on PCrit<sub>global</sub> revealed no significant difference in perceived criticism between groups (hypothesis 1b), $F(1, 102) = 1.65, n.s.$ Because the depressed group had significantly lower levels of marital adjustment, $t(103) = 3.65, p < .001$, any between-group differences, or lack thereof, may have been due to differences in marital adjustment. To control for such differences and to test for marital adjustment’s relationship with perceived criticism, a one-way ANCOVA with Group as a between subjects factor, marital adjustment as a covariate, and PCrit<sub>global</sub> as the dependent variable was conducted. In this ANCOVA, the
main effect of Group was not significant, \( F(1, 101) = .31, \) ns and therefore hypothesis 1c, that the depressed spouses would perceive more criticism than non-depressed spouses after controlling for marital adjustment, was not supported. However, marital adjustment was a significant covariate, \( \beta = -.51, F(1, 101) = 30.26, p < .001, \) suggesting that spouses who are martially distressed are more likely to perceive criticism from their spouse.

3.2 Part Two: Laboratory Criticism

Turning to criticism from the laboratory discussion task, participants in both groups perceived and intended less than one mild critical comment each 30-second interval during the desired change discussion (see Table 1). Observed criticism occurred at a rate of less than one sixth of a critical comment per 30-second interval, which would be equivalent to approximately two and a half intervals with a mild critical comment or one interval with moderately critical comments (see Table 1).

To test hypothesis 2a, that there would be an interaction between Group and Perspective, two separate 2 (Group) \( \times \) 2 (Perspective) split plot ANOVAs were conducted in order to avoid a three level Perspective factor which would have included perceived, intended, and observed perspectives. Because our interest was in interaction contrasts specifically, differences between (a) PCrit_{lab} and ICrit_{lab} and (b) PCrit_{lab} and OCrit_{lab} in the depressed and non-depressed groups, we decided to run two separate ANOVAs, one corresponding to (a) the “insider” perspective and one corresponding to (b) the “outsider” perspective. The first ANOVA used PCrit_{lab} and ICrit_{lab} as the Perspectives. Contrary to hypothesis 2a, there was no significant interaction between
Group and Perspective, $F(1, 95) = .18, ns$. Furthermore there was not a significant main effect of Group, $F(1, 95) = 1.48, ns$ or main effect of Perspective, $F(1, 95) = .33, ns$.

The second ANOVA used PCrit\text{lab} and OCrit\text{lab} as the Perspectives. Like the previous ANOVA and contrary to hypothesis 2a, there was not a significant interaction between Group and Perspective, $F(1, 93) = 2.94, ns$ nor was there a main effect of Group, $F(1, 93) = 1.46, ns$. However, there was a main effect of Perspective, $F(1, 93) = 69.14, p < .001$ with scores on PCrit\text{lab} being higher than those for OCrit\text{lab}.

In sum, taking the mean of the 20 ratings from each 30-second interval during the spousal discussion, depressed participants tended to perceive and be subjected to similar amounts of criticism as did non-depressed participants. What is more, depressed spouses reports of criticism were not any more dissimilar from either intended or observed criticism than were such reports of perceived criticism in non-depressed spouses. It appears that in both groups, there was correspondence between what was perceived and intended but there was less correspondence with outside observer ratings in that outside judges tended to rate less criticism overall than did spouses.

Severity of criticism. In the previous $2 \times 2$ ANOVAs, the dependent variable was the mean of criticism from the twenty intervals during the discussion. Recalling that criticism on each interval was rated on a 0-3 scale, it is important to note that mean scores across the twenty intervals combine both frequency and severity of criticism. Therefore, in further analyses we broke criticism down by level (e.g., by ratings of “1”, “2” and “3”) to separate any effects of criticism intensity from criticism frequency in the twenty intervals. The means and standard deviations of the different levels of criticism for the depressed and control groups are presented in Table 2. For both depressed and control
TABLE 2
MEANS (STANDARD DEVIATIONS) OF LABORATORY CRITICISM BROKEN DOWN BY INTENSITY LEVEL

<table>
<thead>
<tr>
<th>Level of Criticism</th>
<th>Perceived Control $M (SD)$</th>
<th>Perceived Depressed $M (SD)$</th>
<th>Intended Control $M (SD)$</th>
<th>Intended Depressed $M (SD)$</th>
<th>Observed Control $M (SD)$</th>
<th>Observed Depressed $M (SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.22 (.23)</td>
<td>.21 (.15)</td>
<td>.26 (.22)</td>
<td>.23 (.17)</td>
<td>.11 (.14)</td>
<td>.09 (.11)</td>
</tr>
<tr>
<td>2</td>
<td>.07 (.11)</td>
<td>.13 (.14)</td>
<td>.10 (.16)</td>
<td>.14 (.20)</td>
<td>.01 (.05)</td>
<td>.02 (.04)</td>
</tr>
<tr>
<td>3</td>
<td>.07 (.20)</td>
<td>.09 (.14)</td>
<td>.07 (.21)</td>
<td>.07 (.20)</td>
<td>.00 (00)</td>
<td>.00 (00)</td>
</tr>
</tbody>
</table>

NOTE: Perceived, intended, and observed criticism refer to ratings made by target spouses, source spouses, and outside observers, respectively.

couples, $PCrit_{lab}$ of level 1 (that is, criticism of short duration and low intensity) was present about 20% of the time, whereas perceived criticism of levels 2 and 3, (longer duration and higher intensity criticisms) were more rare, occurring in 7-13 percent of the intervals during the interaction.

t-tests were conducted to test for differences among groups on the different intensities (e.g., ratings of “1”, “2”, and “3”) of criticism. For these analyses, significant effects would complicate the interpretation of the signal detection analyses that are presented in the subsequent part of this section. Therefore, we did not adopt a family-wise alpha level. There was only one significant effect that would have been eliminated had we adopted a family-wise alpha level, mainly the significant difference that emerged between groups for perceived criticisms of level 2, $t(106) = -2.32, p < .05$. Therefore, it
appears that there is no evidence that depressed spouses are subjected to a different intensity of criticism than non-depressed spouses.

*Valence of criticism.* The means and standard deviations of the different types of criticism based on valence, (e.g., constructive, neutral, or destructive criticism) are presented in Table 3 along with the overall valence of the entire conversation. Recalling that valence was coded “1” for positive, “2” for neutral, and “3” for negative, both spouses and outside observers rated the overall valence (valence summed across the 20 intervals) as neutral, that is, approximately 2.0. Similarly, for all reporters, instances of neutral criticism were most frequent. It is interesting to note that it appears that spouses in both groups tended to perceive less constructive criticism and more destructive criticism than spouses reported intending. Finally, outside observers tended to indicate low levels of both constructive and destructive criticism, with these occurring on average only 2 – 5 % of the intervals.

While there are many possible ways to explore the means above, we conducted *t*-tests to explore differences between groups on constructive and destructive criticism in order to test the main hypotheses of interest regarding the valence of criticism, or specifically, whether (2c) depressed people perceive more destructive criticism and (2d) perceive less constructive criticism than controls. Results failed to support hypothesis 2c that depressed spouses would perceive more destructive criticism than controls, *t*(96) = -1.34, *ns* and hypothesis 2d, that depressed spouses would perceive less constructive criticism than controls, *t*(96) = .41, *ns*. There were also no significant differences between groups on perceived neutral criticism, *t*(96) = -1.16, *ns*. Similarly, there were no significant differences between groups on any of the types of criticism (destructive,
### TABLE 3
MEANS (STANDARD DEVIATIONS) OF TYPE OF CRITICISM AND OVERALL VALENCE DURING THE LABORATORY DISCUSSION

<table>
<thead>
<tr>
<th>Measure</th>
<th>Perceived</th>
<th></th>
<th>Intended</th>
<th></th>
<th>Observed</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Depressed</td>
<td>Control</td>
<td>Depressed</td>
<td>Control</td>
<td>Depressed</td>
</tr>
<tr>
<td></td>
<td>$M \ (SD)$</td>
<td>$M \ (SD)$</td>
<td>$M \ (SD)$</td>
<td>$M \ (SD)$</td>
<td>$M \ (SD)$</td>
<td>$M \ (SD)$</td>
</tr>
<tr>
<td>Constructive criticism</td>
<td>.08 (.17)</td>
<td>.07 (.11)</td>
<td>.13 (.23)</td>
<td>.12 (.21)</td>
<td>.03 (.08)</td>
<td>.02 (.07)</td>
</tr>
<tr>
<td>Neutral criticism</td>
<td>.17 (.21)</td>
<td>.22 (.18)</td>
<td>.15 (.18)</td>
<td>.18 (.17)</td>
<td>.12 (.12)</td>
<td>.13 (.11)</td>
</tr>
<tr>
<td>Destructive criticism</td>
<td>.13 (.19)</td>
<td>.19 (.25)</td>
<td>.12 (.18)</td>
<td>.15 (.20)</td>
<td>.05 (.15)</td>
<td>.03 (.07)</td>
</tr>
<tr>
<td>Overall valence</td>
<td>1.74 (.51)</td>
<td>1.81 (.56)</td>
<td>1.70 (.50)</td>
<td>1.77 (.53)</td>
<td>1.94 (.33)</td>
<td>1.97 (.27)</td>
</tr>
</tbody>
</table>

NOTE: Perceived, intended, and observed criticism and valence refer to ratings made by target spouses, source spouses, and outside observers, respectively.

neutral, or constructive) for reports of intended or observed criticism. Once again it appears that there is no evidence that depressed spouses are either subject to or perceive different types of criticism as non-depressed spouses.

3.3 Part Three: Signal Detection Analyses of Discussion Data

To further clarify the nature of the relationships among perceived, intended and observed criticism, we subjected the discussion data to signal detection analysis to determine sensitivity and bias in participants’ perceptions of criticism. Separate signal
detection analyses were conducted to compare (a) perceived and intended criticism (P-I) and (b) perceived and observed criticism (P-O).

Recalling that each 10-minute discussion was rated in 30-second intervals, to calculate signal detection scores, each pair of scores (P-I, P-O) was compared at each of the 20 time intervals to determine hit (H) and false alarm (F) rates of each target participant. Because of our interest in sensitivity and bias of perceived criticism, for P-I and P-O comparisons, I and O served as “signals” to be detected or not. If couples had missing data for more than five out of the 20 intervals, or if they had either no criticism or else criticism on every interval, they were excluded from the signal detection analyses because false alarm or hit rates could not be calculated or would be based on an insufficient number of observations. As a result, for the final analyses, there were 80 (59 control and 21 depressed) and 73 (52 control and 21 depressed) participants with P-I and P-O scores, respectively. Couples who were excluded did not differ systematically on any study variables from those who were retained. As shown in Table 4, on average, couples were detecting criticism slightly more than half the time it was present (H) and were erroneously reporting criticism (F) a little over one third of the time. Likewise, H and F were higher when perceived criticism was compared to outside observers than when it was compared to their spouses’ reports of intended criticism. Finally, in each case, couples in the depressed group had slightly higher rates of H and F.

After hit and false alarm rates were computed, we calculated bias and sensitivity. Ordinarily with a graded response (e.g., 0-3) measure one would use additive ROC procedures (Macmillan & Creelman, 2005, p. 126-130), but in our data very few couples
TABLE 4
MEANS (STANDARD DEVIATIONS) OF SIGNAL DETECTION INDICES FOR
OVERALL CRITICISM

<table>
<thead>
<tr>
<th>Measure</th>
<th>P-I Control M (SD)</th>
<th>P-I Depressed M (SD)</th>
<th>P-O Control M (SD)</th>
<th>P-O Depressed M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>.54 (.76)</td>
<td>.60 (.67)</td>
<td>.53 (.68)</td>
<td>.57 (.78)</td>
</tr>
<tr>
<td>Bias</td>
<td>.34 (.79)</td>
<td>-.06 (.91)</td>
<td>.22 (.86)</td>
<td>-.24 (.79)</td>
</tr>
<tr>
<td>Hit Rate</td>
<td>.47 (.26)</td>
<td>.62 (.27)</td>
<td>.51 (.26)</td>
<td>.69 (.22)</td>
</tr>
<tr>
<td>False Alarm Rate</td>
<td>.31 (.28)</td>
<td>.39 (.31)</td>
<td>.34 (.29)</td>
<td>.46 (.30)</td>
</tr>
</tbody>
</table>

NOTE: P-I treats intended criticism as the criterion against which perceived criticism is indexed. P-O treats observed criticism as the criterion against which perceived criticism is indexed.

used all four response options (usually neglecting the rating of “3”). In order to preserve the sample size, we collapsed responses into the more familiar binary form, comparing zero to non-zero (viz., “1”, “2”, and “3”) ratings. This recoding then represents the simple presence versus absence of criticism during each 30-second interval.¹

¹ Obscured by the binary recoding is the possibility that some spouses systematically express criticism in more obvious ways than do others, making the task easier for their partners. If data are converted to binary format, and if depressed or maritally distressed participants are systematically exposed to more obviously critical comments, then these participants might report more criticism, not because they are exposed to more or perceive more, but simply because their task is easier. Therefore, before committing to the binary representation, we tested for severity differences between groups as was presented in section 3.2. Similarly, severity levels were correlated with marital adjustment. Results indicated that neither depression nor marital adjustment was related to the numbers of intended or observed criticism ratings of “1”, “2”, or “3.”
Although there are several parametric binary sensitivity and bias indices to choose from, to index sensitivity we chose \( d' \), \( d' = z(H) - z(F) \), where \( z \) is the standard normal deviate, \( H \) is the hit rate, and \( F \) is the false alarm rate (Macmillan & Creelman, 2005), and we indexed bias using \( c \), where \( c = -0.5 \left[ z(H) + z(F) \right] \), and \( z \), \( H \), and \( F \) are defined as above. Both \( d' \) and \( c \) were chosen because they are traditional signal detection indices recommended by several researchers (Macmillan & Creelman, 2005; Pastore, Crawley, Berens, & Skelly, 2003). Bias and sensitivity calculations produce infinite values for perfect hit and false alarm rates. Therefore, as suggested by Macmillan & Creelman (2005, p. 8), we converted perfect proportions of 0 and 1 to \( 1/(2N) \) and \( 1-1/(2N) \), where \( N \) = the number of trials in the proportion.

**Signal detection of criticism.** The means and standard deviations of sensitivity and bias for P-I and P-O analyses are presented in Table 4. As shown, participants in both groups were not very accurate/sensitive in distinguishing between critical and non-critical comments, with mean scores for \( d' \) ranging from .53 to .60 out of a maximum possible perfectly accurate score of 3.29 (viz., a hit rate of 1 and a false alarm rate of 0). Means of bias scores reveal an interesting pattern. In both cases (P-I and P-O), the depressed group had a negative bias (-.06 and -.24 for \( c_{P-I} \) and \( c_{P-O} \), respectively) whereas the control group had a positive bias (.34 and .22 for \( c_{P-I} \) and \( c_{P-O} \)). The magnitude of this bias can be judged by considering that the range of possible scores for bias on this task was from -1.64 to 1.64, with zero indicating an equal tendency to falsely miss or incorrectly identify criticism.

As might be expected, given the means described above, there was not a significant difference in sensitivity between groups for either \( d'_{P-I}, t(78) = 1.92, p > .05 \) or
$d'_{P-O}, t(71) = .23, p > .05$ and therefore hypothesis 3a was not supported. However, for bias, differences in $c_{P-I}$ scores approached significance, $t(78) = -32, p = .06$, and the difference in $c_{P-O}$ scores was significant, $t(78) = 2.13, p < .05$, supporting hypothesis 3b. Thus, depressed participants seemed to be just as accurate as controls, although both groups had a low level of accuracy. In contrast, depressed spouses tended to be significantly more biased towards reporting criticism than did non-depressed spouses.

Since differences in bias scores may have been due to differences in marital adjustment between the depressed and control groups, we performed two one-way ANCOVAs with marital adjustment as a covariate and either $c_{P-I}$ or $c_{P-O}$ as the dependent variable. There were no significant effects for the $c_{P-I}$ ANCOVA. For the $c_{P-O}$ ANCOVA, the differences between groups on $c_{P-O}$ became non-significant, $F(1, 70) = 1.57, ns$ and marital adjustment emerged as a significant covariate, $\beta = .30, t(70) = 2.52, p < .05$. Therefore, marital adjustment seems to account for the association between depression and bias and thus hypothesis 3d was not supported.

Since both the hit rate ($H$) and false alarm rate ($F$) both contribute equally to bias, we further investigated bias to see if depressed people have a relative tendency, compared to controls, to see criticism when none exists ($F$) and/or to evidence a relative tendency to correctly identify criticism when it does exist ($H$). $t$-tests revealed that for both reports (P-I and P-O) there was a difference in hit rates between depressed and control groups, $t(84) = -2.28, p < .05, t(71) = -2.81, p < .01$ for $H_{P-I}$ and $H_{P-O}$, respectively. There were no significant differences between groups for the false alarm rates.

Again, since differences (or lack thereof) could have been due to differences in marital adjustment, we conducted ANCOVAs with marital adjustment as a covariate and
either hit or false alarm rate as the dependent variable. These ANCOVAs revealed that the only significant effects emerged when perceived criticism was judged in comparison to outside observers. For, $H_{P,O}$ the main effect of Group remained significant, $F(1, 70) = 4.54, p < .05$, whereas marital adjustment was not a significant covariate, $\beta = -.20, t(70) = -1.70, p = .09$. Interestingly, for $F_{P,O}$, marital adjustment emerged as a significant covariate, $\beta = -.31, t(87) = -3.00, p < .01$, whereas the main effect of Group was non-significant, $F(1, 87) = 1.04, ns$.

In sum both depression and marital discord are related to negatively biased reports of criticism. Marital adjustment seems to account for depression’s relationship with bias; however, it does not account for depression’s association with the ability to correctly identify criticism when it is present.

Signal detection of destructive criticism. The above analyses considered all types of criticism (destructive, neutral, and constructive criticism). However, what may be more illuminating is investigating sensitivity and bias with respect to reports of destructive criticism. Only 52 (39 control and 13 depressed) couples had instances of intended destructive criticism and only 22 couples (15 control and 7 depressed) had at least one instance of destructive criticism as rated by outside observers. Furthermore, when it did occur, rates of destructive criticism in couples were generally low. Thus, due to the infrequency of signals and the exclusion of many couples from analyses, these tests were not conducted. Instead, instances of negative valence (despite whether they were tied with critical comments or not) were analyzed in order to increase sample size and power.
Valence signal detection. As with instances of destructive criticism, there were very few couples who used negative valence (either as rated by the source partner or outside raters) during the 10-minute intervals. Therefore, the sample size in the remaining tests was dramatically reduced in comparison to the size for criticism analyses. Regardless, in order to provide preliminary tests of hypotheses 2e and 2f, we performed t-tests comparing groups on sensitivity and bias for detection of negative valence for both P-I and P-O reports. None of these tests showed significant differences between depressed and control groups.

3.4 Part Four: Relationships between Criticism Assessments.

To further understand the meaning of what spouses’ are considering when giving ratings of PCrit_{global}, we conducted correlational analyses of the perceived criticism item with other criticism constructs in this study. Specifically, we tested the assumption that reports of PCrit_{global} would be related to actual criticism (as reported by the source spouse or outside observer) during a laboratory discussion (hypothesis 4a). Results suggest that in the non-depressed sample, PCrit_{global} was related to specific instances of criticism in a laboratory discussion (see Table 5) as reported by the spouses themselves (PCrit_{lab}), the spouses intending the criticism (ICrit_{lab}), or outside observers (OCrit_{lab}). However, in the depressed sample, reports of PCrit_{global} were not significantly related to criticism in a laboratory discussion for any of the reporters even though there was evidence of mild association with PCrit_{lab} and even a modestly negative association with ICrit_{lab}.

Further exploratory analyses tested the PCrit_{global} measure collapsing across diagnostic groups (see Table 6). These analyses lent some support for hypothesis 4b that
TABLE 5

ZERO-ORDER CORRELATIONS BETWEEN PCrit\textsubscript{global} AND OTHER CRITICISM MEASURES WITHIN GROUPS.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Correlation with PCrit\textsubscript{global}</th>
<th>Control</th>
<th>Depressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICrit\textsubscript{global}</td>
<td>.47***</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>PCrit\textsubscript{lab}</td>
<td>.32**</td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td>ICrit\textsubscript{lab}</td>
<td>.22</td>
<td>-.24</td>
<td></td>
</tr>
<tr>
<td>OCrit\textsubscript{lab}</td>
<td>.32**</td>
<td>.01</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: PCrit\textsubscript{global} and ICrit\textsubscript{global} refer to reports of perceived and intended global criticism on the self-report questionnaire. PCrit\textsubscript{lab}, ICrit\textsubscript{lab}, and OCrit\textsubscript{lab} refer to perceived, intended, and observed criticism in the laboratory discussion. Except for ICrit\textsubscript{global} and ICrit\textsubscript{lab}, which are reported by source spouses and Ocrit\textsubscript{lab}, which is reported by outside raters, all other variables are reported by target spouses.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Bias in the laboratory task would be related to global assessments of criticism, as PCrit\textsubscript{global} was related to $c\textsubscript{P-O}$, $r(72) = -.31$, $p < .01$ and the correlation with $c\textsubscript{P-I}$ approached significance, $r(78) = -.22$, $p = .06$. Besides testing this hypothesis, PCrit\textsubscript{global} was also tested for its relation to other laboratory criticism assessments. Specifically, PCrit\textsubscript{global} was related to perceived destructive criticism, $r(96) = .32$, $p < .001$ and level 3 perceived criticisms (severe and characterological criticisms), $r(104) = .25$, $p < .05$.

PCrit\textsubscript{global} was not however related to constructive or neutral criticism or less severe
### Table 6
**Zero-Order Correlations Between PCrit_{global} and Other Perceived Criticism Measures Across Groups**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Correlation with PCrit_{global}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1 Criticism</td>
<td>.17</td>
</tr>
<tr>
<td>Level 2 Criticism</td>
<td>.09</td>
</tr>
<tr>
<td>Level 3 Criticism</td>
<td>.25*</td>
</tr>
<tr>
<td>Destructive Criticism</td>
<td>.32***</td>
</tr>
<tr>
<td>Neutral Criticism</td>
<td>.17</td>
</tr>
<tr>
<td>Constructive Criticism</td>
<td>-.03</td>
</tr>
<tr>
<td>Sensitivity (P-I/P-O)</td>
<td>-.03 / -.19</td>
</tr>
<tr>
<td>Bias (P-I/P-O)</td>
<td>-.22 / -.31**</td>
</tr>
<tr>
<td>Hit Rate (P-I/P-O)</td>
<td>.17 / .22</td>
</tr>
<tr>
<td>False Alarm Rate (P-I/P-O)</td>
<td>.24* / .31**</td>
</tr>
</tbody>
</table>

**Note:** PCrit_{global} refers to perceived global criticism reported by the target spouse on the self-report questionnaire. Levels (1-3) and types (destructive, neutral, and constructive) of criticism presented in this table refer to reports of criticism from target spouses. P-I treats intended criticism as the criterion against which perceived criticism is indexed. P-O treats observed criticism as the criterion against which perceived criticism is indexed.

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

Critical comments. Furthermore, it appears that PCrit_{global} is unrelated to the accuracy of detecting critical comments but is related to the tendency to report criticism. Specifically, PCrit_{global} relates to false alarms, or reporting criticism when none exists. Therefore, those who report more perceived criticism in general, also report criticism when none exists.
4.1 Depression, Marital Adjustment, and Criticality Bias

The main purpose of this study was to examine the possibility that in comparison to those who are not depressed, depressed people over-report instances of spousal criticism. We investigated this question using signal detection methods to compare reports of perceived criticism with criticism as rated by one’s spouse and outside observers on a clip-by-clip basis during a replay of a videotaped spousal discussion. Results suggested that both depressed and less maritally adjusted spouses had a negative reporting bias in that they tended to err on the side of over-reporting criticism rather than missing it. On the other hand, non-depressed participants had a positive reporting bias in that they tended to under-report criticism rather than reporting it when it did not exist. These findings are consistent with the operation of self- and relationship-destructive cognitive biases in depression and discord that increase the perceived criticality of one’s partner and self- and relationship-enhancing biases in non-depressed and maritaly adjusted spouses that reduce the perceived criticality of one’s partner.

These findings are reminiscent of Beck’s (1967) early suggestion that the negative cognitive biases observed in depressed patients are counterparts to positive cognitive biases observed in non-depressed people. Indeed, depression researchers have shown
support, albeit sometimes inconsistent support, for these two kinds of biases. For example, it has been shown that non-depressed people have the tendency to be positively biased in that they attend to more positive environmental stimuli such as positively valenced words (Gotlib, McLachlan & Katz, 1998; McCabe & Gotlib, 1995) whereas depressed people have negative information processing biases and attend to negative environmental stimuli such as negatively valenced words and facial expressions of sadness (Gotlib et al., 2004; see Gotlib & Kraspernova, 1998 for a review). Similar cognitive patterns have been found in children at high risk of developing depression in comparison to those at low risk (Joormann, Talbot, & Gotlib, 2007), suggesting that such factors may even place one at risk for developing depression.

Extending work on biases in depression, social psychologists have studied such biases in interpersonal relationships with their work on the relationship-enhancing effects of positive perceptual distortions (e.g., Murray, Holmes, & Griffin, 1996a; Simpson, Gangestad, & Lerma, 1990). This line of research is similar to the effects found here in that spouses who are satisfied with their marriages perceive more positive characteristics in their spouses than do either the spouses themselves or outside observers. Those in less satisfying marriages demonstrate the opposite effect and view their spouse as more negative than may actually be the case (Murray, Holmes, Dolderman, & Griffin, 2000).

Therefore, evidence from these two lines of research and our finding that both depression and discord were related to criticality bias, suggests that criticality bias may reflect a latent global cognitive style that may underlie relations between depression and marital discord. Importantly, interspousal criticality bias may be primarily linked to marital discord rather than having unique associations with depression as depression was
no longer related to bias once differences in marital adjustment were taken into account. While our measure captured information about cognitions regarding one’s self (e.g., listening for negative comments about the self) and about one’s spouse (e.g., judging one’s spouse as critical or negative), our measure of bias was gleaned from a marital interaction task, so it is perhaps not surprising that it was more closely tied to broader relationship functioning than it was to depression. In other words, the desired change task used in this study was focused primarily on judging one’s spouse as critical. Therefore criticality bias may primarily capture attributions and feelings about one’s spouse and thus be more strongly tied to marital adjustment. Attributions and feelings about oneself may have captured cognitions uniquely related to depression. Perhaps if the task had been to rate one’s own critical behaviors, depressed spouses would have had more of a criticality bias and may have judged themselves as more critical or more negative than they actually were. In sum, it may be that interspousal criticality bias is part of a global cognitive style related to depression because it captures information about one’s view of oneself as well as depressed people’s general negative view of the world. However, interspousal criticality bias may be primarily linked to marital discord because of the unique tie it has with spousal cognitions.

Criticality bias may have been related primarily to relationship functioning rather than depression in this study for several other reasons that fit with current thinking on cognitive vulnerability in depression (see Just, Abramson, & Alloy, 2001 for a review). For example, negative mood activation occurring in maritaly distressed spouses during the 10 minute interaction may have primed negative cognitions in the cognitively vulnerable in the control group thus reducing differences between groups (Just et al.,
Likewise, it is possible that only when depressed people’s schemas are activated via marital discord do they demonstrate such cognitive biases (Beck, Rush, Shaw, & Emery, 1979). In essence, marital discord could have primed either mood or maladaptive schemas. Consistent with vulnerability-stress models of depression (e.g., Abramson, Metalsky, & Alloy, 1989; Beck, 1967), it may be that only in the presence of discord do these cognitive biases emerge in the depressed group. Such biases may also emerge in the non-depressed group as a direct result of discord in spouses at high risk of developing depression, or those with a cognitive vulnerability to depression. It would be interesting to test whether criticality bias places discordant spouses at risk for developing future depression in the face of marital discord.

Despite such findings that marital adjustment was uniquely related to overall criticality bias, the reader is reminded that the above bias reflects the overall tendency to report one’s spouse as critical regardless of whether comments were critical or not. Inspecting the components that make up this bias, some specific effects were found for depression. While depression was not related to the tendency to report criticism when it did not exist (the false alarm rate), it was uniquely related to an increased tendency to correctly identify criticism when it did exist. This finding remained significant even when controlling for marital adjustment. Therefore, it is likely that depressives do not so much as perceive more criticism than exists but rather are more likely to perceive it when it does exist. This is similar to previous work on “depressive realism” (Alloy & Abramson, 1988) and with studies showing that depressed people tend to correctly identify more negative stimuli, such as their own negative behaviors (Roth & Rehm, 1980) or facial expressions (Harkness, Sabbagh, Jacobson, Chowdrey, & Chen, 2005). It may have been
because critical comments imply dislike of negative behaviors that depressed spouses were more readily able to identify such comments which were, after all, in line with their own self-schemas.

While depression was uniquely related to correctly identifying critical comments, marital adjustment was uniquely related to perceiving criticism when none exists. This finding is similar to those on negative relationship attributions that predict future marital discord (Fincham & Bradbury, 1993). Perhaps because they misread partner intentions, make negative attributions for partners’ behaviors, and/or simply judge one’s partner rather than listening to what he or she actually has to say, martially discordant spouses are vulnerable to describing one’s spouse as critical even when in reality their spouse does not intend or act critical.

In sum, that both depression and marital discord are related to bias suggests that interspousal criticality bias may represent an underlying global cognitive style common to both depression and discord that may reflect how one views of oneself and one’s spouse. Differential relations between depression, discord, and the factors that make up bias (e.g., $H$ and $F$) suggest that criticality bias is capturing distinct but related phenomena such as enhanced reporting of actual criticism ($H$) versus reporting criticism when none exists ($F$). Future studies are needed to clarify both cross-sectional and longitudinal relations between depression, discord, and bias. Finally, that sensitivity was unrelated to both depression and discord and that bias, $H$, and $F$ were associated with depression and discord to varying degrees, points to the important distinction between sensitivity ($d'$), bias ($c$), correctly identifying information ($H$), and falsely identifying information when none exists ($F$). In light of such differential correlates for all four
variables (e.g., \( d', c, H, \) and \( F \)), future researchers are urged to be mindful of the distinctions between “accuracy” (the difference between hits and false alarms), “bias” (hits and false alarms), “hits” (correctly identifying information when present), and “false alarms” (falsely identifying something as present when it is not) because treating these as the same construct may create discrepant results and confusion in the literature.

4.2 Global Perceived Criticism

The second purpose of the study was to measure and compare global reports of perceived criticism with criticism coded from a marital interaction task. Previous studies have compared perceived and intended criticism without knowledge of actual criticism (Hooley & Teasdale, 1989) or else previous investigators have assessed levels of actual spousal criticism from interviews in the patients’ absence, assuming behavior in the interview reflects the way relatives actually interact with the patient (Hooley, et al., 1986; Vaughn & Leff, 1976a). Our results with regards to the assumption that perceived and expressed criticism are related are mixed. In the non-depressed group, results support the notion that global reports of perceived criticism are associated with expressed criticism during a specific laboratory task either as rated by spouses or outside observer. However, in the depressed sample, global perceived criticism was not related to any of the laboratory ratings of criticism.

Recalling that depressed spouses showed bias even in a laboratory task when they were watching a videotape of their discussion, it is possible that depressed spouses have even more bias when rating global perceived criticism than they did during the laboratory task. Indeed, the global self-report measure may tap into depressed people’s negative
memory biases (Coyne & Gotlib, 1983; Gotlib & Krasnoperova, 1998) as well as their information processing biases. If depressed spouses do have a large bias in global ratings of perceived criticism, it could account for the non-significant relationship between global ratings and laboratory criticism as rated by spouses and outside observers. Even though we found that mean levels of spousal reports of global perceived and intended criticism were generally similar, and thus perhaps suggest that bias was not at play, it is important to note that the similarity in levels of perceived and intended criticism from the laboratory did not capture the criticality bias that was present because participant’s negative and positive biases cancel each other out in mean analyses. Therefore, it is quite possible that spouses demonstrated bias in global perceived criticism despite the fact that on average, spouses perceived and intended global criticism at similar levels.

Alternatively, spouses may have been responding to the laboratory setting and may have been monitoring how negative they were towards their spouse. Likewise, the way our interaction task was designed may not have appropriately elicited criticism from the source spouses. For example, since the topic of the discussion was something about the depressed target spouse, depressed spouses may have talked throughout most of the discussion, making criticism from the source less likely during these specific 10 minutes.

It is interesting to note however, that several studies of depressed samples have found mixed support for convergence between observed criticism in the lab and global perceptions of criticism (see Hooley & Parker, 2006 for a review). For example, in a sample of remitted depressives, global perceived criticism was modestly, but not significantly correlated with criticism from EE assessments in the laboratory ($r = .27$; Hooley & Teasdale, 1989). In a sample of outpatients with anxiety disorders (Chambless,
Bryan, Aiken, Steketee, & Hooley, 1999) however there was no correlation ($r = .09$) between global perceptions of criticism and CFI coded hostility. Therefore, it appears that the global perceived criticism question may be measuring something different in different samples and that at this time it can not be regarded as synonymous with “actual” criticism or with criticism as measured by outside observers, especially in clinical samples.

Indeed, further supporting the notion that global perceived criticism can not be equated with criticism as measured by outside observers, we also found that global reports of perceived criticism were related to criticality bias in the interaction task, specifically to the tendency to report criticism when it does not exist. Therefore, as suggested by Hooley and Teasdale (1989), reports of perceived criticism appear to reflect how much criticism is “getting through” to spouses and that often times more criticism is “getting through” or being remembered by spouses than actually existed in the case of those spouses who report their spouses as extremely critical. As such, not only do reports of perceived criticism reflect the actual criticism directed at the perceiver, but they also reflect the perceiver’s cognitive bias towards over- or under-perceiving criticism. It is perhaps because global perceived criticism reflects both criticality bias and actual criticism, that it is such a strong predictor of response to treatment (Chambless & Steketee, 1999) and relapse from depression (Hooley & Teasdale, 1989), relative to “actual” observed criticism as measured, say, by the Camberwell Family Interview or Five Minute Speech Sample, which do not capture criticality bias. If so, criticality bias might be a helpful additional focus of interventions for depression and marital discord, beyond the more traditional goal of reducing actual criticism through communication and
problem-solving training (e.g., Beach et al., 1990; Jacobson, Dobson, Fruzzetti, Schmaling, & Salusky, 1991; O’Leary & Beach, 1990).

4.3 Types of Criticism

The third purpose of the study was to differentiate between types of criticism and to explore their correlations with depression and global reports of perceived criticism. In regards to depression, we did not find evidence that depressed spouses are exposed to more criticism overall or that such criticism is in some way different (e.g., more destructive or more intense) than the criticism that non-depressed spouses are exposed to. This is largely in contrast to findings of many studies which have found that in comparison to controls, couples with depressed spouses display more negative and less positive behaviors during marital interactions (e.g., Gotlib & Whiffen, 1989; Johnson & Jacob, 1997).

Finally, in regards to global perceptions of criticism, it appears that high levels of perceived criticism are linked to more severe and more destructive forms of criticism. Spouses do not necessarily consider more constructive or positive forms of criticism when they are globally thinking about spousal criticism. Other research involving non-depressed undergraduate samples generally supports these findings in that only destructive, as opposed to constructive criticism, was related to how globally critical undergraduates rate significant people in their lives (Renshaw, 2006). Therefore, it seems that reports of global criticism correspond with traditional definitions of criticism as defined by EE researchers (e.g., Brown & Rutter, 1966; Rutter & Brown, 1966; Vaughn & Leff, 1976b). This makes the finding that global perceived criticism is minimally
related to such assessments even more illuminating in that even though it appears that spouses and observers share a similar definition of criticism, their reports often differ. Thus, this further supports the notion that criticality bias or cognitive processes are reflected in the perceived criticism construct.

4.4 Limitations and Future Directions

This study employed novel measures of criticism and novel methods of data analysis, so some caution is appropriate when interpreting its findings. It is reassuring to note, however global and laboratory perceived and intended criticism converged in their associations with depression (e.g., none of the overall criticism measures were associated with depression). Likewise, that laboratory intended criticism was significantly associated with laboratory perceived criticism reported by one’s spouse further supports the reliability and validity of these measures, as does their significant associations with observed criticism ($r_s = .29 - .40$). Finally, laboratory interaction measures of perceived, intended, and observed criticism correlated quite highly with their global counterparts, at least in the non-depressed sample.

It is also important to note that in most cases, coders of videos were not completely blind to diagnostic group. Some of the coders were those who had conducted the diagnostic interview with couples whereas in other cases coders may have inferred which spouses were depressed on the basis of gender of the focus spouse and depressive behaviors during the interaction. However, such measures of observed criticism were generally correlated with spouses’ own reports of criticism. Therefore, it does not appear
that observers were influenced in their criticism ratings based on which groups couples were in.

Caution is due when interpreting non-significant effects given the low power and small sample size of the depressed group, especially in the signal detection portion of the analyses. It would be incorrect to rule out any non-significant findings as evidence that such effects do not exist. Future replications of this work with a larger sample need to be undertaken before more definite conclusions can be made. Future researchers might also profitably deploy this paradigm to determine the specificity of over-perception bias across diagnostic groups (e.g., anxiety, schizophrenia) and across communication features (e.g., anger, demandingness).

Additionally, future researchers might investigate gender differences; such differences were not tested in this study due to the fact that there were only four male depressed spouses. Given that women twice as likely as men to experience a depressive episode (Nolen-Hoeksema, 1987), that association between depressive symptoms and marital adjustment seems to be significantly greater for women than men (Whisman, 2001), and that studies have suggested there are important sex differences in the perception of marital communication (Noller, 1980; Carels & Baucom, 1999), gender differences may be important to test.

This study focused on both mean ratings of the laboratory criticism across the 20 intervals and signal detection indicies which are a composite of information from all 20 intervals. Such aggregation across the intervals might obscure differences in patterns of criticism during the discussion. In future studies, it might be interesting to investigate patterns of criticism in depressed and non-depressed couples to see if spouses of
depressed partners are merely responding to criticism from the depressed spouse or if depressed spouses respond to partner criticism with counter criticism.

Neither the causal sequence nor the causal mechanisms connecting criticality bias and actual criticism are presently known. It is fairly easy to imagine how prior experience with a critical spouse could heighten the expectation of future criticality, producing a subsequent criticality bias (Gotlib & Whiffen, 1989). It is less easy to imagine the reverse causal direction, that is, to imagine how prior criticality bias may produce later actual partner criticality. Nevertheless, there is evidence that “positive illusions” are associated with later reduced conflict in some couples (Murray, Holmes, & Griffin, 1996b), and this might be a complement process to the one we investigated, which was a kind of “negative illusion.” Therefore, causal paths between criticality bias and actual criticality are possible in both directions, and it will require experimental studies to determine the relative contributions of each path.

The longitudinal course of criticality bias in depression should be investigated to see if it is a vulnerability factor predicting depression, or if it is only activated during depressive states or when negative mood is induced via negative marital interactions. Likewise, future work might investigate if negative criticality bias can explain longitudinal associations between depression and marital discord or if might serve as an individual difference factor heightening the depressive effects of the experience of marital discord.
<table>
<thead>
<tr>
<th>Clip Start Time</th>
<th>Describe Your Thoughts and/or Feelings</th>
<th>Was your partner being positive, negative, or neutral towards you?</th>
<th>How critical was your partner being of you?</th>
<th>Describe Your Partner's Thoughts and/or Feelings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
</tr>
<tr>
<td></td>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>2 = moderately critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = very critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
</tr>
<tr>
<td></td>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>2 = moderately critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = very critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
</tr>
<tr>
<td></td>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>2 = moderately critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = very critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
</tr>
<tr>
<td></td>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>2 = moderately critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = very critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
</tr>
<tr>
<td></td>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.</td>
<td>2 = moderately critical</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = very critical</td>
<td></td>
</tr>
</tbody>
</table>
### APPENDIX B: SOURCE SPOUSE VIDEO RATING FORM

<table>
<thead>
<tr>
<th>Clip Start Time</th>
<th>Describe Your Thoughts and/or Feelings</th>
<th>Were you being positive, negative, or neutral towards your partner?</th>
<th>How critical were you intending to be of your partner?</th>
<th>Describe Your Partner’s Thoughts and/or Feelings</th>
</tr>
</thead>
<tbody>
<tr>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
<td></td>
</tr>
<tr>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2 = moderately critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>3 = very critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
<td></td>
</tr>
<tr>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2 = moderately critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>3 = very critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
<td></td>
</tr>
<tr>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2 = moderately critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>3 = very critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td>I was thinking and/or feeling:</td>
<td>+</td>
<td>0 = not at all critical</td>
<td>Partner was thinking and/or feeling:</td>
<td></td>
</tr>
<tr>
<td>because:</td>
<td>0</td>
<td>1 = slightly critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2 = moderately critical</td>
<td>because:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>3 = very critical</td>
<td>because:</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C: DYADIC ADJUSTMENT SCALE (SPANIER, 1976)

University of Notre Dame Marital Therapy and Research Clinic

Dyadic Adjustment Scale (Spanier, 1976)

<table>
<thead>
<tr>
<th>Always Agree</th>
<th>Almost Agree</th>
<th>Somewhat Disagree</th>
<th>Frequency Disagree</th>
<th>Almost Disagree</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Handling family finances.</td>
<td>2</td>
<td>Matters of recreation.</td>
<td>3</td>
<td>Religious matters.</td>
</tr>
<tr>
<td>6</td>
<td>Sex relations.</td>
<td>7</td>
<td>Conventionality (correct or proper behavior).</td>
<td>8</td>
<td>Philosophy of life.</td>
</tr>
<tr>
<td>12</td>
<td>Making major decisions.</td>
<td>13</td>
<td>Household tasks.</td>
<td>14</td>
<td>Leisure-time interests and activities.</td>
</tr>
</tbody>
</table>

16. How often do you discuss or have you considered divorce, separation, or terminating your relationship?

17. How often do you or your mate leave the house after a fight?

18. In general, how often do you think that things between you and your partner are going well?

19. Do you confide in your mate?

20. Do you ever regret that you married? (Or lived together)

21. How often do you and your partner quarrel?

22. How often do you and your mate "get on each other’s nerves”?

23. Do you kiss your mate?

24. Do you and your mate engage in outside interests together?

25. How often would you say the following events occur between you and your mate?

26. Have a stimulating exchange of ideas.

27. Laugh together.

28. Calmly discuss something.

29. Work together on a project.

30. These are some things about which couples sometimes agree and sometimes disagree. Indicate if each item below caused differences of opinion or were problems in your relationship during the past few weeks (check yes or no).

31. The numbers on the following indicate different degrees of happiness in your relationship. The middle point, "happy," represents the degree of happiness of most relationships. Please fill in the circle that best describes the degree of happiness, all things considered, of your relationship.

32. Fill in a circle—Extremely Unhappy Fairly Unhappy A little Unhappy Happy Very Happy Extremely Happy

33. Which of the following statements best describes how you feel about the future of your relationship?

34. My relationship can never succeed, and there is no more that I can do to keep the relationship going.
APPENDIX D: GENERAL CRITICISM SCALE

(COUTURE, 2001; COUTURE ET AL., 2000; SMITH ET AL. 2001)

General Criticism Scale

Please use the 0 - 5 scales provided to rate how much criticism you believe generally characterizes each of the situations described.

Thank you for reading each item carefully before circling the best number.

Generally speaking... 

1. How critical is your spouse of you?
   Not at all Critical 1 2 3 4 Extremely Critical 5

2. How critical are you of your spouse?
   Not at all Critical 1 2 3 4 Extremely Critical 5

3. How critical are people, other than your spouse, of you?
   Not at all Critical 1 2 3 4 Extremely Critical 5

4. How critical are people, other than you, of your spouse?
   Not at all Critical 1 2 3 4 Extremely Critical 5

5. How self-critical are you?
   Not at all Critical 1 2 3 4 Extremely Critical 5

6. How self-critical is your spouse?
   Not at all Critical 1 2 3 4 Extremely Critical 5

7. How critical is your spouse of people other than you?
   Not at all Critical 1 2 3 4 Extremely Critical 5

8. How critical are you of people other than your spouse?
   Not at all Critical 1 2 3 4 Extremely Critical 5
REFERENCES


