The Effect of Stress on Empathic Accuracy in Romantic Couples

Alexander O. Crenshaw, Karena Leo, and Brian R. W. Baucom
University of Utah

Accurately understanding the thoughts and feelings of romantic partners, termed empathic accuracy, is critical for optimal relationship functioning. Empathic failure is linked to common reasons couples seek therapy (Doss, Simpson, & Christensen, 2004; Jacobson & Christensen, 1996) and is either implicitly or explicitly a target of many couple therapies (e.g., Jacobson & Christensen, 1996). More specifically, couple therapies target partners' abilities to accurately understand one another preceding and during conflict, periods characterized by high levels of stress. The current study tests the hypothesis that acute stress can be harmful for empathic accuracy in romantic couples, and tests two competing path models of the impact of stress on accuracy. Results show that an acute stressor affected accuracy of men and women differently, impairing accuracy in women but not observably affecting men's accuracy. The effect of the stressor on empathic accuracy for women was mediated by curvilinear arousal, and men's accuracy was also associated with curvilinear arousal. This pattern of results suggests that moderate arousal is optimal for empathic accuracy for both men and women, but this effect was twice as large for women relative to men. These findings point toward potential avenues to improve existing couple therapies by incorporating strategies to mitigate the effects of stress, thereby increasing couples' ability to accurately understand one another in therapy. These findings also suggest strategies for couples to be optimally productive when having important relationship conversations by attending to their level of arousal.

Keywords: empathic accuracy, stress, conflict, marriage, arousal

Supplemental materials: http://dx.doi.org/10.1037/fam0000508.supp

Accurately understanding the thoughts and feelings of romantic partners, termed empathic accuracy (Ickes, 1997), is critical for optimal relationship functioning. For example, couples entering therapy most frequently present with concerns about communication difficulty and lack of emotional affection (Doss et al., 2004), problems linked to poor or inaccurate understanding one another's thoughts, feelings, and needs (Kilpatrick, Bissonnette, & Rusbult, 2002; Jacobson & Christensen, 1996; Verhofstadt et al., 2016).

Additionally, improving empathic functioning is an implicit or explicit target of many empirically supported couple therapies (Epstein & Baucom, 2002; Jacobson & Christensen, 1996; Johnson & Denton, 2002). However, being empathically accurate is difficult, even under ideal circumstances and among established romantic partners who have known each other for years (e.g., Verhofstadt et al., 2016). The difficulty of empathic accuracy can be tied to the complex and ambiguous nature of empathic inference, as both empathy theories and empirical research suggest it is a finely tuned and sensitive process (Decety & Jackson, 2004; Preston & de Waal, 2002). These factors make empathic accuracy potentially susceptible to disruption by outside factors, such as stress (e.g., Arnsten, 2009; Decety & Jackson, 2004; Kudielka, Hellhammer, Kirschbaum, Harmon-Jones, & Winkielman, 2007). Relationship problems co-occur with greater life stressors (e.g., Conger, Rueter, & Elder, 1999) and are associated with increased vulnerability to existing stressors (e.g., Karney & Bradbury, 1995). Consequently, couples seeking therapy, a situation under which empathic accuracy is especially important, may have particular difficulty understanding one another.

It is therefore important to understand more about situational factors, such as stress, that make empathic inference more or less difficult. Doing so can point toward ways to improve existing interventions by implementing strategies to optimize partner's ability to accurately understand each other's thoughts and feelings. Stress is a common life experience with wide-reaching effects on cognitive, emotional, and behavioral processes (Arnsten, 2009; Kudielka et al., 2007), making it likely to substantially impact empathic accuracy.

Moreover, stress is responsive to various forms of intervention, which points toward clear ways to improve empathic accuracy by addressing stress. As a first step toward understanding the connection between stress and empathic accuracy, the present study tests the effect of an experimentally induced stressor on empathic accuracy, and additionally proposes a path model by which stress and empathic accuracy are linked: subjective emotional arousal.

Empathic Accuracy in Romantic Relationships

A recent meta-analysis links greater empathic accuracy with higher levels of relationship satisfaction for both members of a couple (Sened et al., 2017a). Likewise, empathic accuracy appears to contribute to a host of relationship-promoting behaviors (for a more in-depth discussion, see Baucom & Atkins, 2013). For example, greater empathic accuracy is associated with more skillful provision of social support (Verhofstadt et al., 2016), more accommodative behaviors during conflict (Kilpatrick et al., 2002), and, when partners are motivated to be responsive, more responsive behavior (Winczewski, Bowen, & Collins, 2016).

Tuning into the thoughts and emotional states of another, and making accurate inferences about those states, is difficult even under ideal circumstances: accuracy of empathic inferences among romantic partners typically falls at or below 30% (e.g., Verhofstadt et al., 2016; Winczewski et al., 2016). This low level of accuracy is consistent with the complex nature of accurately inferring the internal state of a partner, which requires substantial commitment of cognitive and emotional resources (Bird & Viding, 2014; Decety & Jackson, 2004). Partners must tune into verbal content, nonverbal vocal and visual information, the person's behavior, knowledge about the situation and the person's past response to similar situations, and so forth, and loss of any information tends to reduce empathic accuracy (e.g., Hall & Schmid Mast, 2007). The difficulty, complexity, and often-ambiguous nature of empathic inference suggests empathy is a finely tuned process (e.g., Decety & Jackson, 2004), making it sensitive to even minor disruptions.

Stress has far-reaching effects on cognitive, emotional, and behavioral processes (Arnsten, 2009; Kudielka et al., 2007), and the sensitive nature of empathic accuracy makes it potentially vulnerable to these effects. For example, modern models of empathy emphasize the importance of regulatory systems (e.g., emotion regulation and self-other distinction) for optimal empathic functioning (Decety & Jackson, 2004; Preston & de Waal, 2002). These systems rely on structures in the prefrontal cortex (Decety & Jackson, 2004), which are sensitive to stress (see Arnsten, 2009, for a review). Stress also impacts cognitive functioning in ways directly related to empathic inference. For example, stress increases the likelihood of erroneously attributing behavior to personality characteristics without adequately considering situational factors, as well as increasing the negativity of attributions made toward others (Kubota et al., 2014).

This collection of findings suggests that stress impairs empathic accuracy such that the lowest levels of empathic accuracy occur under the highest levels of stress. If this hypothesis is correct, it could reflect either a linear or a curvilinear association between empathic accuracy and stress. The possibility of a linear associations between stress and empathic accuracy is the most intuitive: as stress increases, empathic accuracy decreases. However, studies of

arousal and performance suggest that the association between stress and empathic accuracy may be curvilinear (e.g., Yerkes & Dodson, 1908). Whereas high levels of stress should result in high levels of arousal that impair cognitive performance, moderate amounts of stress should create moderate levels of arousal, which are optimal for sustaining motivation, attention, and cognitive functioning (e.g., Lupien, Maheu, Tu, Fiocco, & Schramek, 2007). It is therefore possible that while high levels of stress are associated with low levels of empathic accuracy, moderate levels of stress may be associated with higher levels of empathic accuracy.

Testing both possibilities is vital for clarifying the potential clinical implications of stress on empathic accuracy. Many relationally dissatisfied couples who seek couple therapy display high levels of acute arousal during relationship conflict (e.g., Baucom et al., 2015), but there appears to be a subset of couples who display the opposite problem of underarousal during conflict. For example, female partners who exhibit very low levels of arousal during conflict are at increased risk for failing to benefit from behaviorally based couple therapies (e.g., Baucom, Atkins, Simpson, & Christensen, 2009). Results consistent with the linear hypothesis would suggest that efforts to enhance empathic accuracy based on arousal should solely focus on lowering arousal when very high, whereas results consistent with the curvilinear hypothesis would suggest such efforts should focus both on lowering arousal when it is very high and increasing it when it is very low.

Additionally, acute stress affects men and women in different ways depending on the nature and duration of the eliciting event, suggesting possible sex differences in the effect of stress on empathic accuracy. For example, women tend to report more subjective distress from social stressors compared with men (e.g., Kelly, Tyrka, Anderson, Price, & Carpenter, 2008) and show greater sensitivity to ACTH, which stimulates the secretion of the stress hormones glucocorticoids (Kudielka, Buske-Kirschbaum, Hellhammer, & Kirschbaum, 2004), whereas men exhibit larger cortisol and ACTH responses to social stressors (Kudielka et al., 2007). One study also found stress to increases egocentricity in men while reducing it in women (Tomova, von Dawans, Heinrichs, Silani, & Lamm, 2014). Taken together, these findings suggest that stress may impact empathic accuracy differently for men and women, but the evidence is contradictory regarding the specific direction of this difference.

The Current Study

The current study tests whether acute stress reduces empathic accuracy by testing the effect of a stress manipulation on empathic accuracy in a sample of committed dating and married couples, and tests linear and curvilinear associations between arousal and empathic accuracy. We hypothesize that couples randomly assigned to undergo an experimentally induced stressor will show lower levels of empathic accuracy toward one another compared with couples who are not stressed (Hypotheses 1). We also hypothesize that arousal and empathic accuracy will be associated in a curvilinear fashion such that moderate levels of arousal will be associated with the highest levels of empathic accuracy and both very low and very high levels of arousal will be associated with the lowest levels of empathic accuracy (Hypothesis 2). We test a path model explaining associations between stress, linear arousal, curvilinear arousal, and empathic accuracy, and predict that curvilinear arousal will mediate the association between stress and em-

pathic accuracy (Hypothesis 3). Finally, we examine moderation of these associations by sex in an exploratory fashion due to lack of consistent empirical evidence regarding empathic accuracy-relevant sex differences in stress response. To access study materials, including procedures manuals and measures, see (https://osf.io/k3nq2/).

Method

Participants

Participants were 48 couples ($N = 96$ individuals)[1] ages 18–50 recruited from a university subject pool, campus and community flyers, and online advertisements. We used a stratified sampling design to both maximize inclusion and ensure variability in relationship functioning by requiring that at least 20 couples fall above and below the estimated population mean of 16 on the Couples Satisfaction Index-4 (CSI-4; Funk & Rogge, 2007) during screening, according to the lowest CSI-4 score of the couple. Past research (as well as a recent meta-analysis; Sened et al., 2017a) suggests empathic accuracy is most important early in relationships (Kilpatrick et al., 2002), so recruitment was limited to those in early, committed relationships in order to test associations in the population for which empathic accuracy is most important. Participants were required to either be married for less than two years or monogamously dating for at least one year and currently cohabiting. These requirements resulted in similar self-reported relationship length between married and unmarried couples ($B = -1.58$ months dating, $p = .854$).

Participants self-identified as White (76%), Asian (11.5%), Native Hawaiian/Pacific Islander (2.1%), American Indian or Alaska Native (1%), or declined to answer (9.4%). Eleven participants (11.5%) were Hispanic or Latino. Two couples were same-sex. Thirty-six couples were married, while 12 were unmarried. Married couples were married for an average of 11.5 months ($SD = 7.02$), while the average total relationship length in the sample was 39.4 months ($SD = 25.3$). Median individual annual income was $18,600 ($SD = $17,586$). Twenty-seven (28.1%) participants had a high school diploma or equivalent, 53 (55.2%) had a college degree, 15 (15.6%) had a graduate or professional degree, and one (1%) declined to answer. Average self-reported relationship satisfaction was 65.7 ($SD = 14.5$) on the Couples Satisfaction Index-16 (Funk & Rogge, 2007). Ten couples were in the clinically distressed range (CSI-16 < 51.5).

Procedure

The present study is part of a larger study involving a 3-hr procedure that included collection of physiological measures, physiological baseline tasks, a battery of self-report questionnaires, one conflict discussion, and a video recall procedure outlined below. All procedures were approved by the University of Utah Institutional Review Board (Protocol number: IRB_00078872; Title: Expression and perception during marital communication), and participants provided consent upon arrival to the lab. This article focuses on several self-report questionnaires, the conflict discussion, and verbal thought/feeling ratings from the video recall procedures.

Participants completed a battery of self-report questionnaires followed by the Problem Areas Questionnaire (PAQ; Heavey,

Christensen, & Malamuth, 1995), a measure of common areas of conflict in romantic relationships, to identify a topic for the conflict discussion. For the video-taped discussion, the experimenter chose the topic with the highest cumulative disagreement across both partners and which both partners agreed to discuss. Participants were then instructed to discuss the topic with one another for 10 min and to try to understand the issue as best as possible after the experimenter left the room.

Next, participants completed a brief postdiscussion questionnaire and were then asked to individually complete a video recall procedure to measure empathic accuracy (Ickes, Stinson, Bissonnette, & Garcia, 1990). For the first review of the video (self-rating 1), partners separately watched a recording of the conflict discussion and were prompted every 60 seconds to write down any specific thought or feeling from the past minute of the conversation.[2] Participants were instructed to only record thoughts or feelings they distinctly recalled having during the conversation, not what they think or feel when watching the video.

Prior to the second review of the video, participants completed one of two randomly assigned conditions (see online supplemental material for a discussion of task order within study procedures). In the stress condition, partners separately completed the Trier Social Stress Test (TSST; Kudielka et al., 2007), a standardized and well-validated social stressor involving a mock job interview and speeded mental arithmetic. In the control condition, participants rated nature pictures for an equivalent length of time. Participants then completed the second step of the video recall procedure (partner rating); during this step, participants watched the conflict discussion video again and recorded specific thoughts or feelings they believed their partner had during each minute. Finally, to examine consistency in participants' responses and rule out a possible alternative explanation for Hypothesis 1, participants repeated the procedure a third time, in which they recorded their own thoughts and feelings from the conversation a second time (self-rating 2). Participants were compensated their choice of either $35 or (for students) 3 hr of research credit.

Measures

Empathic accuracy. Empathic accuracy was calculated using the Ickes method (Ickes et al., 1990), which operationalizes em-

---

[1] Sample size determination: We planned to collect 50 couples in total, based on an effect size from a study using a similar manipulation where $d = 1.08$ (Ambady & Gray, 2002), which was adjusted downward to be conservative and to account for the nested data structure. After adjusting to account for data dependency, a sample size of 50 couples (100 individuals) was estimated to have .8 power to detect effect sizes of $d = .63$. One couple was dropped from analysis due to not following instructions, and another couple dropped out just before the designated end date of data collection, leaving 48 couples.

[2] This procedure is modeled around that developed by Ickes and colleagues (e.g., Ickes et al., 1990), with one modification. The original procedure allows participants to stop the recording themselves any time they recall having had a thought or feeling. The procedure was adjusted for the current study by using software to automatically stop the recording every one minute, rather than allowing participants to decide, in order to allow for comparison of the two self-ratings. According to Ickes (W. Ickes, personal communication, August 15, 2014), such a modification is unlikely to adversely affect the validity of the video recall paradigm. Additionally, a recent meta-analysis suggests the procedures are comparable (Sened et al., 2017a).

pathic accuracy as the match between a person's inferences about their partner's thoughts and feelings with the partner's self-ratings of thoughts and feelings. Six trained research assistants, blind to condition, compared corresponding self and partner thought/feeling entries at each minute and rated them on a three-point scale (2: essentially the same content; 1: similar but not the same content; 0: essentially different content). The coding team met weekly with the first author to review the past week's codes and discuss discrepancies toward ensuring reliability of future codes. Two coders failed to achieve acceptable agreement with the others and were dropped from analyses. Scores for the entire interaction were summed and averaged across the four remaining coders, then divided by the total possible score to get a final empathic accuracy score for each partner ranging from 0 to 1. Cronbach's alpha among coders was .899, consistent with past studies (e.g., Ickes et al., 1990; Winczewski et al., 2016).

Self-rating test-retest score. We computed a self-rating test-retest score using the same 3-point coding system to score similarity of each partner's self-rating completed before the stress manipulation with the one completed after. Doing so measured participants' consistency in self-reported thoughts and feelings before and after the stress manipulation using the same metric as empathic accuracy (0–1). Cronbach's alpha among coders was .936.

Valence and arousal. Participants completed the Self-Assessment Manikin (SAM; Bradley & Lang, 1994), a well-validated visual analogue measure, before and immediately after the stress or control task to measure emotional valence and arousal. The SAM has three continuous dimensions (valence, arousal, and dominance), and participants are asked to rate how they are feeling on each dimension. Higher scores represent higher levels of arousal and more positive valence. Dominance was not considered for this study.

Social sensitivity. Previous research suggests that individuals vary in ability to detect the emotions in others (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). To minimize error variance due to individual variation in trait empathic ability, participants completed the Reading the Mind in the Eyes Test—Revised (Eyes Test; Baron-Cohen et al., 2001) as part of the initial battery of questionnaires, which was a planned covariate for use in all analyses (we also report results of analyses excluding Eyes Test in the online supplemental material, which are extremely similar in terms of direction, magnitude, and significance of results). On the Eyes Test, participants identify the emotion shown in 36 pairs of eyes from among four choices. This measure shows good ability to detect individual differences in social sensitivity (Baron-Cohen et al., 2001). Cronbach's alpha was .576, consistent with past studies (Vellante et al., 2013).

Analytic Plan

Analyses were conducted using multilevel models (MLM) estimated using HLM 7.01 (Raudenbush, Bryk, & Congdon, 2011); robust standard errors for predictors in these models are reported. Hypothesis 1 used the following multilevel model (presented in series of equations format) to examine the effect of stress condition on empathic accuracy, controlling for individual differences in social sensitivity and exploring the moderating influence of sex:

Level-1:

$$\text{EMPATHIC ACCURACY}_{ij} = \beta_{0j} + \beta_{1j} \times (\text{SEX})$$
$$+ \beta_{2j} \times (\text{SOCIAL SENSITIVITY}) + r_{ij}$$

Level-2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \times (\text{CONDITION}) + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \times (\text{CONDITION})$$

$$\beta_{2j} = \gamma_{20}$$

where i represents partners, j represents couples, SEX is effect-coded (male = $-.5$; female = $.5$), SOCIAL SENSITIVITY is grand-centered scores on the Eyes Test, and CONDITION is dummy coded stress condition (0 = control; 1 = stress).

Hypotheses 2 (curvilinear association between arousal and empathic accuracy) was tested by regressing empathic accuracy onto sex, arousal, arousal squared, the interaction between sex and arousal, the interaction between sex and arousal squared, and social sensitivity (covariate). Significant interactions with sex were decomposed into simple slopes to clarify interpretation. A significant effect of arousal squared provides evidence for the curvilinear hypothesis, whereas, if the effect of arousal squared is nonsignificant, the model would be rerun with only linear effects to test the linear hypothesis. Indirect paths from condition to empathic accuracy (Hypothesis 3) were tested using the RMediation web tool (Tofighi & MacKinnon, 2011). Arousal and social sensitivity were grand-centered as predictors prior to creation of interaction terms and analysis.

Results

Preliminary Analyses

Table 1 presents means, standard deviations, and zero-order correlations for study variables (see online supplemental material for separate tables by condition and sex). Contrary to expectations, empathic accuracy was not significantly correlated with the Eyes Test or relationship satisfaction, and relationship length was not significantly associated with empathic accuracy across conditions (however, see Sened et al., 2017a for context on associations between empathic accuracy, relationship length, and satisfaction).

Randomization and Manipulation Checks

We tested for randomization failure by testing associations between stress condition and Eyes Test, age, relationship length, and relationship satisfaction. All associations were nonsignificant, indicating randomization was successful. Details of these analyses are available from the first author. By chance, there were identical numbers of married ($n = 18$) and unmarried ($n = 6$) couples in each condition. We performed a manipulation check by regressing valence and arousal ratings onto condition, sex, and sex $\times$ condition. As expected, there was a significant effect of condition on valence ($B = -4.95$, $p < .001$) and arousal ($B = 4.53$, $p < .001$) such that participants in the stress condition reported experiencing more negativity and more arousal than those in the control condition. There was not a significant effect of sex nor a significant sex by condition interaction on either valence or arousal ($ps > .336$),

Table 1
Means, Standard Deviations, and Correlations for Variables, Collapsed Across Sex and Condition

Variable	M (SD)	Correlations						
		1	2	3	4	5	6	7
1. Empathic accuracy (0–1)	.21 (.10)	—						
2. Test-retest (0–1)	.44 (.15)	.074	—					
3. Social sensitivity (0–36)	27.60 (3.8)	.139	−.023	—				
4. Age	25.45 (3.9)	.002	−.170	−.020	—			
5. Relationship length	39.40 (25.3)	−.127	−.012	.207*	.106	—		
6. Satisfaction (0–81)	65.71 (14.5)	.074	.067	.025	−.084	.080	—	
7. Valence (0–16.2)	6.93 (4.3)	.194†	.072	.022	.028	−.009	.000	—
8. Arousal (0–16.2)	7.39 (4.4)	−.103	−.148	.016	−.090	.058	.180†	−.297**

Note. Test-retest = Test retest score of self-ratings; Relationship length is in units of months; Satisfaction = relationship satisfaction (Funk & Rogge, 2007). Means and correlations are based on raw data prior to winsorizing relationship length.
† $p < .1$.   * $p < .05$.   ** $p < .01$.

indicating that the stress manipulation was successful, and men and women did not significantly differ in amount of stress experienced.

Stress and Empathic Accuracy

Table 2 presents results of the MLM used to test the effect of stress on empathic accuracy (Hypothesis 1).[3] There was a significant conditional main effect of sex ($B = .048$, $p = .032$), such that women were more accurate than men in the control condition. This result was qualified by a significant Sex × Condition interaction ($B = −.079$, $p = .013$), shown in Figure 1. Decomposition of the interaction indicated that men's accuracy did not significantly differ by condition ($B = −.007$, 95% CI [−.066, .051], $p > .500$), while women's did ($B = −.086$, 95% CI [−.138, −.034], $p = .002$). Additionally, although women showed greater accuracy in the control condition, men's and women's empathic accuracy did not significantly differ in the stress condition ($B = −.031$, $p =$

Table 2
Results of MLM Predicting Empathic Accuracy and Self-Rating Test-Retest Scores

Variable	ES	B (SE)	95% CI	p
Model 1: Predicting empathic accuracy (H1)				
Intercept		.231 (.018)	[.196, .266]	<.001
Sex	.46[a]	.048 (.021)	[.006, .090]	.032
Eyes Test	.14[b]	.004 (.002)	[−.001, .009]	.122
Condition	−.45[a]	−.047 (.023)	[−.093, −.000]	.055
Sex × Condition	−.68[c]	−.079 (.031)	[−.139, −.019]	.013
Model 2: Predicting test-retest score				
Intercept		.450 (.021)	[.408, .491]	<.001
Sex	−.34[a]	−.052 (.041)	[−.133, .029]	.218
Eyes Test	−.03[b]	−.001 (.004)	[−.010, .007]	.791
Condition	−.13[a]	−.019 (.031)	[−.079, .041]	.538
Sex × Condition	.08[c]	.024 (.059)	[−.091, .139]	.685

Note. ES = Effect size; SE = Standard error; CI = Confidence interval. Effect sizes should be viewed as estimates given lack of standard method for calculating effect sizes in multilevel models.
[a] Cohen's *d*.   [b] Standardized regression coefficient (i.e., β).   [c] Cohen's *d* using procedures from Selya, Rose, Dierker, Hedeker, and Mermelstein (2012).

.149). These results indicate that women were more empathically accurate than men when not stressed, but exposure to stress significantly reduced their accuracy to that of men, whose accuracy was not significantly affected by the stress induction.

Rather than impacting empathic accuracy specifically, it is possible that stress caused participants to respond more negatively in general. If true, stress should affect participants' second ratings of themselves as well as of their partners. To test this possibility, we regressed self-rating test-retest scores onto gender, condition, and gender × condition. Results of this test (see Table 2) indicate that neither the main effect of condition ($B = −.019$, 95% CI [−.079, .041], $p = .538$), the main effect of sex ($B = −.052$, 95% CI [−.133, .029], $p = .218$), nor the interaction between sex and condition ($B = .024$, 95% CI [−.091, .139], $p = .685$) were significant. These findings indicate that the stress manipulation did not significantly impact how participants completed their self-ratings, suggesting the stress effect is specific to empathic accuracy rather than affecting response style more broadly.

Arousal and Empathic Accuracy

Table 3 present results of the MLMs testing associations between empathic accuracy, arousal, and arousal squared, as well as path models. In testing Hypothesis 2, there was a significant interaction between sex and arousal squared ($B = −.002$, $p = .015$). Shown in Figure 2, decomposition of this interaction revealed a significant effect of arousal squared on empathic accuracy for both men ($B = −.002$, 95% CI [−.003, −.0001], $p = .029$) and women ($B = −.004$, 95% CI [−.005, −.002], $p < .001$). This association between arousal squared and empathic accuracy did not significantly differ by stress condition for either sex (Men: $B = −.0003$, $p = .876$; Women: $B = −.0009$, $p = .643$). These results support the curvilinear arousal-empathic accuracy hypoth-

[3] For a sensitivity analysis, we also added to the model grand-centered age, attachment anxiety and avoidance (Fraley, Waller, & Brennan, 2000), length of relationship, emotional closeness (from the mean of five Likert-scale items such as "I feel close to my spouse"), relationship satisfaction (Funk & Rogge, 2007), and a numerical post-discussion empathic accuracy score. Direction, magnitude, and significance of results remained unchanged with the added variables, except for the conditional main effect of sex, which was reduced to a trend ($B = .043$, $p = .090$). Description of these measures is included in the online supplemental material.
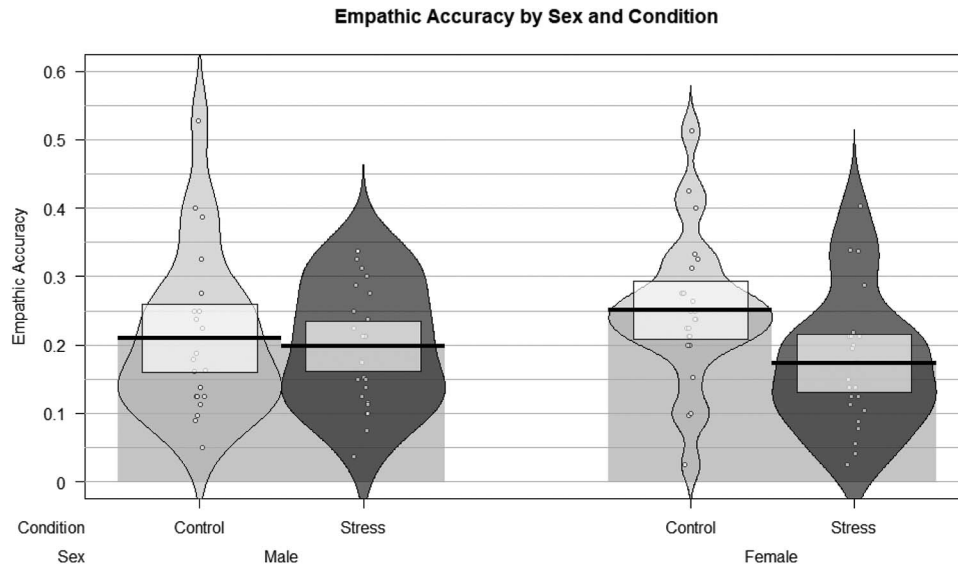
Empathic Accuracy by Sex and Condition

Figure 1. Pirate plots (Phillips, 2017) of empathic accuracy by sex and condition. Black horizontal lines represent group means, boxes around the black lines represent 95% confidence intervals of the mean, dots represent raw data points, and "beans" surrounding the raw data represent distributional properties).

esis (Hypothesis 2) for both men and women, but the association was significantly larger for women.

In the model regressing arousal squared onto sex, condition, social sensitivity, and all lower order terms, there was a significant Sex × Condition interaction in predicting arousal squared ($B = 21.77$, $p < .001$). Decomposing this interaction revealed a significant effect of condition on arousal squared for women ($B = 14.18$, 95% CI [5.94, 22.43], $p = .001$), but not for men ($B = -7.58$, 95% CI [-16.35, 1.19], $p = .086$). Finally, shown in Figure A1 of the supplemental material, there was a significant indirect effect from stress condition, through arousal squared, to empathic accuracy for women ($B = -.045$, 95% CI [$-.085$, $-.015$]), but not for men ($B = .010$, 95% CI [$-.002$, .030]). These results support the curvilinear path model (Hypothesis 3) for women, indicating that curvilinear arousal mediated the association between stress and accuracy, but do not support either the linear or curvilinear path model for men.

Post Hoc Analyses

One possible explanation for the sex differences in the effect of stress is that men and women could have had differences in baseline stress prior to the stress or control task. To test for this possibility, we tested whether men and women differed in arousal measured immediately before the stress or control task (i.e., post self-rating 1) and also reran key analyses while controlling for pretask arousal levels. Pretask arousal levels did not significantly vary by sex ($B = -.263$, $p = .707$). Controlling for pretask arousal, results remained highly consistent in terms of magnitude, direction, and significance, details of which are in the online supplemental material.

Discussion

Improving empathic functioning is of wide interest among couple researchers and therapists, as well as in the social sciences

more broadly. The finely tuned, sensitive, and difficult nature of empathic inference makes it susceptible to outside factors, such as stress. The present study experimentally tested the effect of stress on romantic partners' ability to be empathically accurate with one another and tested two competing models in which stress affects empathic accuracy through either linear or curvilinear arousal. Results indicated that the Trier Social Stress Task (Kudielka et al., 2007) significantly impaired empathic accuracy for women, but not for men. Women were significantly more empathically accurate than men when not stressed, but were negatively affected by the stress induction, resulting in similar accuracy to men. Moreover, results suggest that stress specifically impacted women's empathic abilities rather than impacting how they respond to questions about thoughts and feelings more broadly.

In addition to testing the causal impact of stress, the present study tested two competing models in which the association between stress and empathic accuracy is explained in part by linear or curvilinear arousal. Results supported the curvilinear path model path model for women, indicating that curvilinear arousal mediates the association between stress and empathic accuracy. For men, we also found that curvilinear arousal was associated with accuracy even though the stress manipulation did not significantly impact their accuracy. Finally, the association between curvilinear arousal and empathic accuracy was more than twice as strong for women compared with men, suggesting women's accuracy is more sensitive to arousal than men's. Taken together, this pattern of results indicates that while high levels of stress are associated with significantly lower levels of empathic accuracy only for women, both men and women's accuracy appears sensitive to arousal levels. The curvilinear association between arousal and empathic accuracy for both men and women suggests moderate arousal is optimal and both too little and too much arousal is associated with lower accuracy for both sexes.

These results are consistent with other studies highlighting the importance of arousal for relationship processes. For example,

Table 3
Results of MLMs Testing Associations Between Stress, Arousal, and Empathic Accuracy

| | Hypothesis 2: Testing effect of curvilinear arousal on empathic accuracy | | | | | | | | |
| | Overall | | | | | Men | | Women | |
Variable	ES	B	SE	95% CI	p	B	p	B	p
Intercept		.260	.017	[.226, .293]	<.001				
Sex	.53[a]	.055	.024	[.009, .102]	.023				
Eyes Test	.07[b]	.002	.002	[−.002, .006]	.368				
Arousal	.04[b]	.001	.002	[−.002, .004]	.521				
Arousal^2	−.40[b]	−.003	.001	[−.004, −.002]	<.001	−.002	.029	−.004	<.001
Sex × Arousal	.04[b]	.002	.003	[−.004, .008]	.535				
Sex × Arousal^2	−.64[c]	−.002	.001	[−.004, −.001]	.015				

| | Path model 'a' path | | | | | | | | | |
| | Linear (outcome: arousal) | | | | | Curvilinear (outcome: arousal^2) | | | | |
Variable	ES	B	SE	95% CI	p	ES	B	SE	95% CI	p
Intercept		5.13	.43	[4.29, 5.96]	<.001		17.96	2.02	[14.0, 21.9]	<.001
Sex	.28[a]	1.19	1.22	[−1.21, 3.58]	.336	−.55[a]	−8.92	4.05	[−16.87, −.97]	.033
Eyes Test	.09[b]	−.03	.10	[−.22, .17]	.798	−.16[b]	−.70	.40	[−1.47, .08]	.085
Condition	.17[a]	4.53	.76	[3.05, 6.02]	<.001	.20[a]	3.30	3.26	[−3.08, 9.68]	.316
Sex × Condition	.05[c]	.30	1.66	[−2.96, 3.55]	.860	.88[c]	21.77	5.75	[10.50, 33.03]	<.001

| | Path model 'b' path (Outcome: empathic accuracy) | | | | | | | | | |
| | Linear | | | | | Curvilinear | | | | |
Variable	ES	B	SE	95% CI	p	ES	B	SE	95% CI	p
Intercept		.235	.017	[.200, .269]	<.001		.270	.022	[.227, .313]	<.001
Sex	.51[a]	.053	.021	[.011, .095]	.018	.68[a]	.071	.029	[.014, .127]	.019
Eyes Test	.14[b]	.004	.002	[−.001, .009]	.113	.09[b]	.002	.002	[−.002, .006]	.233
Condition	−.54[a]	−.056	.024	[−.103, −.010]	.022	−.37[a]	−.038	.028	[−.093, .016]	.176
Sex × Condition	.33[b]	−.095	.037	[−.168, −.023]	.014	−.16[b]	−.047	.035	[−.116, .022]	.189
Arousal	.09[b]	.002	.002	[−.002, .006]	.294	.08[b]	.002	.002	[−.001, .005]	.244
Arousal^2	—	—	—	—	—	−.35[b]	−.002	.001	[−.003, −.001]	<.001
Sex × Arousal	.07[c]	.004	.004	[−.005, .012]	.406	.11[b]	.005	.004	[−.003, .013]	.195
Sex × Arousal^2	—	—	—	—	—	−.61[c]	−.002	.001	[−.004, −.000]	.049

Note. ES = Effect size; SE = Standard error; CI = Confidence interval. Effect sizes should be viewed as estimates given lack of standard method for calculating effect sizes in multilevel models.
[a] Estimate of Cohen's *d*. [b] Estimate of standardized regression coefficient (i.e., β). [c] Cohen's *d* from Selya et al. (2012).

Gottman and Levenson (1992) found that couples showing higher arousal during conflict were at greater risk of divorce, which may in part be due to the breakdown of healthy relationship processes under high arousal (Snyder, Simpson, & Hughes, 2006). Arousal may also mean different things for different couples. One study investigating predictors of response to behavioral couple therapies found that low pretreatment arousal predicted better response to treatment in moderately dissatisfied wives, but worse response to treatment for severely dissatisfied wives (Baucom et al., 2009). The authors surmised that low arousal in severely dissatisfied wives may be indicative of disengagement from having already given up on the relationship, consistent with the idea that very low arousal during important interactions may be detrimental for relationship health.

What might account for the detrimental impact of stress in women? It is possible that stress reduced participants' motivation to be empathically accurate. Theorists highlight the central importance of motivation for empathy and empathic accuracy and that motivation can vary across circumstances (Ickes & Simpson, 2008; Zaki, 2014). Empathic accuracy is known to vary by the presence of a particular motivator (Ickes, Gesn, & Graham, 2000) and can be increased by manipulating motivation to be accurate, such as by providing financial rewards (Klein & Hodges, 2001). Sened, Yovel, Bar-Kalifa, Gadassi, and Rafaeli (2017b) found that the occurrence of conflict predicted greater accuracy in daily life for men and suggested that conflict increased men's motivations to be accurate, perhaps due to men's tendency to want to solve conflicts. Simpson and colleagues (see Ickes & Simpson, 2008, for a review) have also found that partners are sometimes motivated to be *in*accurate, particularly when accuracy might be threatening to the relationship. This motivation hypothesis for women could be tested in future research, for example, by crossing a stress manipulation with a motivation-enhancing intervention to determine if increased motivation could overcome the impact of stress.

Additionally, empathy can break down when an individual cannot differentiate the source of an emotion (self vs. other), resulting in personal distress that focuses one's attention inward rather than outward as with empathy (Decety & Lamm, 2009). High levels of arousal in the present study may have had a similar effect by directing partners' attention inward toward managing
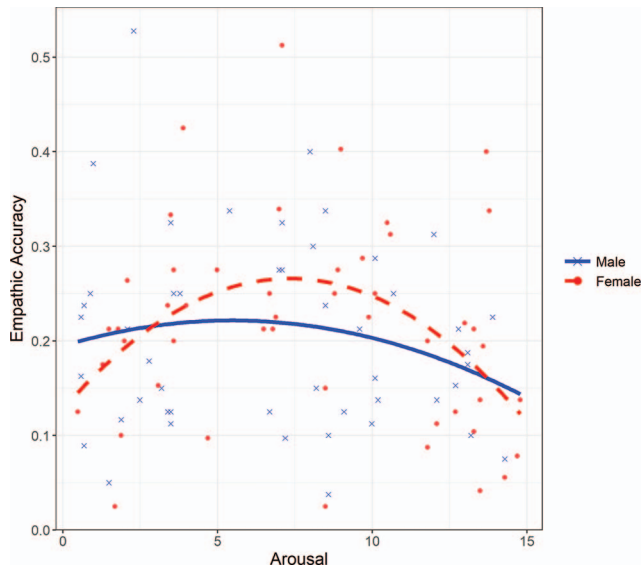
Figure 2. Curvilinear fitted plots of arousal and empathic accuracy for men and women, combined conditions (stress and control). See the online article for the color version of this figure.

their own distress and away from paying attention to the partner (Baucom et al., 2009). Another intriguing possibility is related to stress hormones. Lupien et al. (2007) cite a wealth of evidence for an inverted U association between the stress hormones glucocorticoids and cognitive functioning broadly, similar to the association observed between arousal and empathic accuracy in the present study, suggesting glucocorticoids may biologically mediate associations between stress and empathic accuracy through their association with cognitive functioning.

Empathic Accuracy, Stress, and Sex

The present study found three key sex differences regarding empathic accuracy: (a) women were more accurate than men in the control condition; (b) stress reduced women's empathic accuracy, but did not significantly affect men's; and (c) associations between curvilinear arousal and empathic accuracy were twice as strong for women as men. Some of these differences may result from sex differences in motivation for empathic accuracy. Graham and Ickes (1997) proposed that women are generally more motivated to be empathically accurate due to cultural stereotypes of "women's intuition," for example, particularly in situations when empathy is made salient to participants. Two studies supported the motivational hypothesis for sex differences in empathic accuracy (Ickes et al., 2000; Klein & Hodges, 2001), and one found that sex differences disappear when providing men additional motivation (Klein & Hodges, 2001). In the current study, women may have had greater motivation to be accurate in the control condition, resulting in greater accuracy than men. However, exposure to the acute stress of the TSST may have reduced participants' motivation to be accurate, having a lesser effect on men because their motivation was already lower. This pattern may explain why women in the control condition were more accurate and why stress impaired accuracy for women but had no observable effect on men.

However, under the motivational hypothesis, it is not clear why women's accuracy was more strongly linked to arousal than men's. One possible explanation for this pattern is related to the glucocorticoid hypotheses described earlier. Kudielka et al. (2004) found that young women, compared with young men (our sample was overwhelmingly young), had greater adrenal cortex sensitivity to ACTH, a precursor to glucocorticoids (Kudielka et al., 2004), which could explain the sex difference in arousal sensitivity we observed. However, a separate study found men exhibit larger ATCH responses to social stressors (Kudielka et al., 2007), suggesting the opposite pattern, so it remains unclear why women's accuracy was more strongly linked to arousal.

Another possibility for the sex differences observed is that there may have been a floor effect for men: men, having lower accuracy to begin with, may have not had much room to decrease as a result of the stress manipulation. Ickes et al. (1990) estimated the "floor" for accuracy using the Ickes video recall procedure by pairing dyad members with a randomly selected member of a different dyad and estimating the average "accuracy" of these randomly paired dyads. They found that mean accuracy was 5.8 for randomly paired dyads, compared with 21.7 for true dyads (accuracy was scaled by a factor of 100 compared to the present study). This mean for true dyads was highly similar to the control group means in the present study (Men $M = .21$; Women $M = .25$), suggesting .058 may be a reasonable approximation of floor accuracy in the present sample. This estimate of floor accuracy is 1.3 standard deviations below the male control group mean, suggesting sufficient room to decrease. However, we cannot fully rule out floor effects; future research could test this possibility by identifying if men's accuracy could be reduced through other means.

Test–Retest Reliability of Video Recall Procedure

In addition to examination of empathic accuracy, the current study provides the first estimate to our knowledge of test–retest reliability of the thought/feeling self-ratings used in the Ickes empathic accuracy video recall procedure (e.g., Ickes et al., 1990). The average test-retest score was .44 out of 1, providing a baseline estimate of test–retest reliability (see online supplemental material for a discussion of possible causes of this low reliability). This estimate also provides context for commonly observed low empathic accuracy scores, which typically fall at or below .3 (e.g., Verhofstadt et al., 2016; Winczewski et al., 2016). Whereas a mean empathic accuracy score of .21 is low compared to the maximum *theoretical* score of 1, it is not as low if the maximum *practical* score is closer to .44. This finding suggests couples are substantially better at inferring each other's thoughts and feelings than previous research has suggested.

Clinical Implications

This study is the first to establish stress's role in shaping partner's ability to accurately tune into each other's thoughts and feelings. Regardless of the specific stressor, women are likely to be less empathically accurate toward their partners when highly stressed, and this change is attributable to changes in arousal caused by stress. Moreover, both men and women's empathic accuracy is tied to arousal, in which empathic accuracy is lowest under conditions of very high and very low arousal and is highest

when arousal is moderate. Important conversations may therefore be most productive when both partners are alert and engaged, but not highly stressed.

These findings highlight potential avenues to improve existing couple therapies by attending to partners' arousal level, which could increase partners' ability to accurately understand one another in therapy. For example, therapists might help highly stressed partners better empathize with and understand one another by incorporating arousal-reduction interventions, such as paced breathing or mindfulness practice (e.g., Moqadam & Kazerooni, 2017) into therapy. Conversely, therapists might help underaroused partners through arousal-raising activities. Relatedly, these findings suggest strategies couples may use on their own to be optimally productive when having important relationship conversations. Couples might benefit from knowing that they are likely to have greater difficulty understanding one another when arousal is high (or very low), and may benefit from brief arousal-reduction interventions before conflictual discussions or from delaying conflictual discussions until stress is reduced (e.g., the "time out" intervention in Traditional Behavioral Couple Therapy; Jacobson & Margolin, 1979).

Limitations and Future Directions

There are several limitations to keep in mind when considering these results. First, the sample size was modest, so power to detect smaller effects is limited. For example, although the point estimate for the effect of the stress manipulation on empathic accuracy in men was small ($-.007$, on a 0 to 1 scale), and the 95% confidence interval ($-.066$ to $.051$) is inconsistent with effects larger than a decrease of .066 points, we cannot rule out smaller effects for men. Future research would benefit from investigating these associations with a larger sample. Second, this study tested a causal hypothesis and thus prioritized internal validity over external validity. Future research should examine associations between stress and empathic accuracy across time as they naturally occur to identify the real-world associations between these constructs. Third, the sample was predominantly non-Hispanic White, educated, and heterosexual, and couples were early in their relationship. Future research should test these associations in more diverse samples and with more established couples, especially as empathic accuracy may change over time and may also function differently for more established couples (Bissonnette, Rusbult, & Kilpatrick, 1997; Thomas, Fletcher, & Lange, 1997; also see Sened et al., 2017a). Additionally, future research should use other arousal measures beyond self-report (e.g., psychophysiology). Similarity in the inverted U association between the stress hormones glucocorticoids and cognitive functioning (Lupien et al., 2007) and that observed here between subjective arousal and empathic accuracy is striking, and points toward glucocorticoids as a possible mediator of empathic accuracy, which would be a valuable avenue to pursue in future research.

Conclusion

Empathic accuracy is an important component of healthy relationship functioning. The findings of the current study indicate that high levels of stress are detrimental to empathic accuracy in women, while not noticeably affecting men. It also found that, rather than stress being necessarily harmful, moderate stress appears optimal, while both high and low levels of stress are associated with reduced empathic accuracy for both men and women. These findings suggest that couples, particularly women, will best understand one another, and perhaps in turn navigate important conversations effectively, when alert and engaged but not significantly stressed. Couples therapists may benefit from attending to arousal levels of partners in couple therapy, and incorporating arousal-reduction strategies when partners become heated may help partners understand one another's thoughts and feelings in therapy most effectively.

References

Ambady, N., & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology, 83,* 947–961. http://dx.doi.org/10.1037/0022-3514.83.4.947

Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience, 10,* 410–422. http://dx.doi.org/10.1038/nrn2648

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241–251. http://dx.doi.org/10.1111/1469-7610.00715

Baucom, B. R., & Atkins, D. C. (2013). Understanding marital distress: Polarization process. In M. A. Fine & F. D. Fincham (Eds.), *Handbook of family theories: A content-based approach* (pp. 145–166). New York, NY: Routledge.

Baucom, B. R., Atkins, D. C., Simpson, L. E., & Christensen, A. (2009). Prediction of response to treatment in a randomized clinical trial of couple therapy: A 2-year follow-up. *Journal of Consulting and Clinical Psychology, 77,* 160–173. http://dx.doi.org/10.1037/a0014405

Baucom, B. R., Sheng, E., Christensen, A., Georgiou, P. G., Narayanan, S. S., & Atkins, D. C. (2015). Behaviorally-based couple therapies reduce emotional arousal during couple conflict. *Behaviour Research and Therapy, 72,* 49–55. http://dx.doi.org/10.1016/j.brat.2015.06.015

Bird, G., & Viding, E. (2014). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience and Biobehavioral Reviews, 47,* 520–532. http://dx.doi.org/10.1016/j.neubiorev.2014.09.021

Bissonnette, V. L., Rusbult, C. E., & Kilpatrick, S. D. (1997). Empathic accuracy and marital conflict resolution. In W. Ickes (Ed.), *Empathic accuracy* (pp. 251–281). New York, NY: Guilford Press.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 49–59. http://dx.doi.org/10.1016/0005-7916(94)90063-9

Conger, R. D., Rueter, M. A., & Elder, G. H., Jr. (1999). Couple resilience to economic pressure. *Journal of Personality and Social Psychology, 76,* 54–71. http://dx.doi.org/10.1037/0022-3514.76.1.54

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3,* 71–100. http://dx.doi.org/10.1177/1534582304267187

Decety, J., & Lamm, C. (2009). Empathy versus personal distress: Recent evidence from social neuroscience. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 199–214). Cambridge, MA: MIT Press.

Doss, B. D., Simpson, L. E., & Christensen, A. (2004). Why do couples seek marital therapy? *Professional Psychology: Research and Practice, 35,* 608–614. http://dx.doi.org/10.1037/0735-7028.35.6.608

Epstein, N. B., & Baucom, D. H. (2002). *Enhanced cognitive-behavioral therapy for couples: A contextual approach.* Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10481-000

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

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

Graham, T., & Ickes, W. (1997). When women's intuition isn't greater than men's. In W. J. Ickes (Ed.), *Empathic Accuracy* (pp. 117–143). New York, NY: Guilford Press.

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. *Journal of Personality and Social Psychology, 63,* 221–233. http://dx.doi.org/10.1037/0022-3514.63.2.221

Hall, J. A., & Schmid Mast, M. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion, 7,* 438–446. http://dx.doi.org/10.1037/1528-3542.7.2.438

Heavey, C. L., Christensen, A., & Malamuth, N. M. (1995). The longitudinal impact of demand and withdrawal during marital conflict. *Journal of Consulting and Clinical Psychology, 63,* 797–801. http://dx.doi.org/10.1037/0022-006X.63.5.797

Ickes, W. (Ed.). (1997). *Empathic accuracy.* New York, NY: Guilford Press.

Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships, 7,* 95–109. http://dx.doi.org/10.1111/j.1475-6811.2000.tb00006.x

Ickes, W., & Simpson, J. A. (2008). Motivational aspects of empathic accuracy. In G. J. O. Fletcher & M. S. Clark (Eds.), *Interpersonal processes* (pp. 229–250). *Blackwell handbook of social psychology* Oxford, England: Blackwell.

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology, 59,* 730–742. http://dx.doi.org/10.1037/0022-3514.59.4.730

Jacobson, N. S., & Christensen, A. (1996). *Acceptance and change in couple therapy: A therapist's guide to transforming relationships.* New York, NY: W. W. Norton.

Jacobson, N. S., & Margolin, G. (1979). *Marital therapy: Strategies based on social learning and behavior exchange principles.* New York, NY: Routledge.

Johnson, S. M., & Denton, W. (2002). Emotionally focused couples therapy: Creating secure connections. In A. S. Gurman & N. S. Jacobson (Eds.), *Clinical handbook of couple therapy* (3rd ed., pp. 221–250). New York, NY: Guilford Press.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin, 118,* 3–34. http://dx.doi.org/10.1037/0033-2909.118.1.3

Kelly, M. M., Tyrka, A. R., Anderson, G. M., Price, L. H., & Carpenter, L. L. (2008). Sex differences in emotional and physiological responses to the Trier Social Stress Test. *Journal of Behavior Therapy and Experimental Psychiatry, 39,* 87–98. http://dx.doi.org/10.1016/j.jbtep.2007.02.003

Kilpatrick, S. D., Bissonnette, V. L., & Rusbult, C. E. (2002). Empathic accuracy and accommodative behavior among newly married couples. *Personal Relationships, 9,* 369–393. http://dx.doi.org/10.1111/1475-6811.09402

Klein, K. J. K., & Hodges, S. D. (2001). Gender Differences, Motivation, and Empathic Accuracy: When it Pays to Understand. *Personality and Social Psychology Bulletin, 27,* 720–730. http://dx.doi.org/10.1177/0146167201276007

Kubota, J. T., Mojdehbakhsh, R., Raio, C., Brosch, T., Uleman, J. S., & Phelps, E. A. (2014). Stressing the person: Legal and everyday person attributions under stress. *Biological Psychology, 103,* 117–124. http://dx.doi.org/10.1016/j.biopsycho.2014.07.020

Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., & Kirschbaum, C. (2004). HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: Impact of age and gender. *Psychoneuroendocrinology, 29,* 83–98. http://dx.doi.org/10.1016/S0306-4530(02)00146-4

Kudielka, B. M., Hellhammer, D. H., Kirschbaum, C., Harmon-Jones, E., & Winkielman, P. (2007). Ten years of research with the Trier Social Stress Test—Revisited. In E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior* (pp. 56–83). New York, NY: Guilford Press.

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition, 65,* 209–237. http://dx.doi.org/10.1016/j.bandc.2007.02.007

Moqadam, N. K., & Kazerooni, G. (2017). A study of mindfulness-based stress reduction's effectiveness on marital satisfaction in Women. *International Journal of Health Studies, 3,* 35–38.

Preston, S. D., & de Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25,* 1–20. http://dx.doi.org/10.1017/s0140525x02000018

Phillips, N. (2017). *yarrr: A Companion to the e-Book "YaRrr!: The Pirate's Guide to R"* (R package version 0.1.4). Retrieved from https://CRAN.R-project.org/package=yarrr

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2011). HLM 7 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.

Rusbult, C. E., Verette, J., Whitney, G. A., Slovik, L. F., & Lipkus, I. (1991). Accommodation processes in close relationships: Theory and preliminary empirical evidence. *Journal of Personality and Social Psychology, 60,* 53–78. http://dx.doi.org/10.1037/0022-3514.60.1.53

Sened, H., Lavidor, M., Lazarus, G., Bar-Kalifa, E., Rafaeli, E., & Ickes, W. (2017a). Empathic accuracy and relationship satisfaction: A meta-analytic review. *Journal of Family Psychology, 31,* 742–752. http://dx.doi.org/10.1037/fam0000320

Sened, H., Yovel, I., Bar-Kalifa, E., Gadassi, R., & Rafaeli, E. (2017b). Now you have my attention: Empathic accuracy pathways in couples and the role of conflict. *Emotion, 17,* 155–168. http://dx.doi.org/10.1037/emo0000220

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f2, a measure of local effect size, from PROC MIXED. *Frontiers in psychology, 3,* 111.

Snyder, D. K., Simpson, J. E., & Hughes, J. N. (Eds.). (2006). *Emotion regulation in couples and families: Pathways to dysfunction and health.* Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/11468-000

Thomas, G., Fletcher, G. J. O., & Lange, C. (1997). On-line empathic accuracy in marital interaction. *Journal of Personality and Social Psychology, 72,* 839–850. http://dx.doi.org/10.1037/0022-3514.72.4.839

Tofighi, D., & MacKinnon, D. P. (2011). RMediation: An R package for mediation analysis confidence intervals. *Behavior Research Methods, 43,* 692–700. http://dx.doi.org/10.3758/s13428-011-0076-x

Tomova, L., von Dawans, B., Heinrichs, M., Silani, G., & Lamm, C. (2014). Is stress affecting our ability to tune into others? Evidence for

gender differences in the effects of stress on self-other distinction. *Psychoneuroendocrinology, 43,* 95–104. http://dx.doi.org/10.1016/j .psyneuen.2014.02.006

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry, 18,* 326–354. http://dx.doi.org/10.1080/ 13546805.2012.721728

Verhofstadt, L., Devoldre, I., Buysse, A., Stevens, M., Hinnekens, C., Ickes, W., & Davis, M. (2016). The role of cognitive and affective empathy in spouses' support interactions: An observational study. *PLoS ONE, 11*(2), e0149944. http://dx.doi.org/10.1371/journal.pone .0149944

Winczewski, L. A., Bowen, J. D., & Collins, N. L. (2016). Is empathic accuracy enough to facilitate responsive behavior in dyadic interaction?

Distinguishing ability from motivation. *Psychological Science, 27,* 394– 404. http://dx.doi.org/10.1177/0956797615624491

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology & Psychology, 18,* 459–482. http://dx.doi.org/10.1002/ cne.920180503

Zaki, J. (2014). Empathy: A motivated account. *Psychological Bulletin, 140,* 1608–1647. http://dx.doi.org/10.1037/a0037679