Title

Naturalistic Assessment of Interpersonal Emotion Regulation in Romantic Couples via Ambulatory Acoustic Recording: Feasibility, Reliability, and Validity

Author

Swerdlow, Benjamin Abravanel

Publication Date

2022

Peer reviewed|Thesis/dissertation

Naturalistic Assessment of Interpersonal Emotion Regulation in Romantic Couples via Ambulatory Acoustic Recording: Feasibility, Reliability, and Validity

By

Benjamin Abravanel Swerdlow

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Psychology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Sheri Johnson, Chair
Professor Iris Mauss
Professor Cameron Anderson

Summer 2022

Abstract

Naturalistic Assessment of Interpersonal Emotion Regulation in Romantic Couples via Ambulatory Acoustic Recording: Feasibility, Reliability, and Validity

by

Benjamin Swerdlow

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Sheri L. Johnson, Chair

Recent theoretical and empirical contributions to the emotion regulation and close relationships literatures have begun to elucidate the dynamic complexities and functional consequences of interpersonal emotion regulation (IER) processes. The present study sought to evaluate a naturalistic method for capturing IER processes in the daily lives of committed romantic couples by using ambulatory acoustic recording in tandem with ecological momentary self-report to measure enacted behavior and subjective experience. A coding manual was developed for this purpose. Then, a community sample of cohabiting romantic couples was recruited (final $N = 59$ couples, 118 participants) to consider the feasibility, reliability, and validity of this method. Couples completed a thorough baseline assessment, followed by eight days of ecological momentary assessment and ambulatory acoustic recording with the Electronically Activated Recorder. A team of undergraduate and post-baccalaureate coders was trained to identify and code the content of acoustically observable IER-relevant interactions that occurred within the couples. Participants' self-reported evaluations of wearing the audio recorder, total number of recordings produced, the characteristics of the IER-relevant interactions that coders identified, inter-rater reliability among the coders, the correlational structure of the coder ratings, and correspondences between coder ratings and participant self-report were analyzed. Consistent with past research, participants generally reported that wearing the audio recorder was only modestly obtrusive for themselves and others. In total, participants produced nearly 2200 hours of audio. Using a periodic sampling strategy, coders identified 836 unique IER-relevant interactions. Many of these interactions appeared to have been "invisible" to the participants in that they reported not having engaged in IER during those epochs in ecological momentary self-report surveys. Inter-rater reliability ranged from tolerable to excellent for the coder ratings, albeit some codes were rarely endorsed in this sample. There were systematic correspondences between coder ratings and participant self-report, although the magnitude of these correspondences was generally quite modest. Overall, data from this study suggest that this approach to recording and coding spontaneously occurring IER interactions is feasible, reliable, and valid.

Dedication

To my partner, Vivianne, without whose love and support this dissertation would simply not have been possible.

To my family, who have been on this journey with me from the beginning—literally.

To my lab mates, Andrew, Amy, Devon, Manon, Matt, and my lab twin, Jen, who inspire me endlessly and are the best collaborators a scientist could wish for.

To all of my mentors at Berkeley and especially to my advisor, Sheri Johnson, for their generosity, support, and guidance.

Acknowledgements

The study described in this dissertation would not have been possible without the hard work of a small legion of research assistants, including Alina Yu, Julia Wang, David Baquirin, Sarah Anvar, Vanessa Kim, Robin Stuart, Lauren Price, Mackenzie Maddox, Shirley Yan, Holly Hoover, Allison Hom, Judy Xu, Alicia Luong, Esther Suh, Vivan Nye, Sarah Fleming, Maria Barillas-Hernandez, Kiley Yeaman, Caroline Provost, Jessie Zhao, Jamie Kim, and Megan Wang.

Megan Robbins' consultation and advice on all things Electronically Activated Recorder throughout the inception and early stages of this study was invaluable.

I would also like to extend my gratitude to all of my role models, collaborators, and countless colleagues who have shaped my career, my thinking about interpersonal emotion regulation, and this study in particular, including: Sheri Johnson, Charles Sanislow, Rajita Sinha, Iris Mauss, Lauren Sippel, Lesley Berk, Craig Williams, Jamil Zaki, and Dan Wile.

This study was funded in part by fellowships from the National Institute of Mental Health via the Predoctoral Training Consortium in Affective Science and from the Greater Good Science Center.

Naturalistic Assessment of Interpersonal Emotion Regulation in Romantic Couples via Ambulatory Acoustic Recording: Feasibility, Reliability, and Validity

Emotion regulation has emerged as one of the most widely studied constructs in psychological research over the past twenty-five years. Indeed, tens of thousands of publications about emotion regulation have been authored in the two and a half decades since the enormously generative process model of emotion regulation first debuted (Gross, 1998; Gross, 2013). This strong and sustained interest in emotion regulation can be attributed at least in part to compelling evidence of the functional consequences of emotion regulation, such as the tendency to use specific emotion regulation strategies, across multiple domains of functioning. To wit, past research has linked emotion regulation not only to psychological well-being and mental health, but also to interpersonal functioning, physical health, and even political outcomes (e.g., Aldao, Nolen-Hoeksema, & Schweizer, 2010; Bloch, Haase, & Levenson, 2014; Cloitre et al., 2019; Compas et al., 2017; DeSteno, Gross, & Kubzansky, 2013; Gross, 2002; Halperin, 2014; Kraiss, Klooster, Moskowitz, & Bohlmeijer, 2020; Miu et al., 2022).

In a pioneering study of emotion regulation in everyday life, 98% of instances of emotion regulation recalled to interviewers had occurred in a social (versus non-social) context (Gross, Richards, & John, 2006). This finding was a powerful demonstration that many of the most salient instances of emotion regulation in our everyday lives are embedded in social contexts. This observation of the social embeddedness of emotion regulation was, moreover, consistent with the widely discussed notion—particularly among developmental psychologists—that emotion regulation frequently entails the active involvement of close others, such as a parent or spouse (e.g., Thompson, 1994). Yet, emotion regulation has most typically been studied as an intrapersonal process of *self*-regulation, with the overwhelming majority of emotion regulation studies focusing on non-social emotion regulation (Campos, Walle, Dahl, & Main, 2011). Mindful of this gap, conceptual and empirical advances in the past decade have increasingly directed attention toward interpersonal dimensions of emotion regulation (e.g., Gross, 2015; Netzer, Van Kleef, and Tamir, 2015; Niven, Totterdell, & Holman, 2009; Williams et al., 2018). In part, this increased emphasis on interpersonal dimensions of emotion regulation reflects the view that emotional functioning and social functioning are fundamentally entangled in a reciprocal relationship (cf. Beckes & Coan, 2012; Shaver & Mikulincer, 2014).

The emerging field of interpersonal emotion regulation (IER) focuses on the "slice of interpersonal interactions deliberately devoted to influencing one's own (intrinsic) or others' (extrinsic) emotions" (Dixon-Gordon et al., 2015, p. 37; see also Niven, 2017; Reeck, Ames, & Ochsner, 2016; Zaki & Williams, 2013). Whereas it has long been acknowledged that the mere presence of conspecifics can shape emotional experiences and emotion regulation (e.g., Bratec et al., 2020; Jakobs, Manstead, & Fischer, 1999; Schacter, 1959), IER has been distinguished from these mere presence effects as being a *goal-directed process* whereby two or more individuals engage in live social transactions that are motivated to regulate their own or each other's affective states (Zaki & Williams, 2013). For example, individuals might vent their emotions to a close friend after a frustrating day at work in the hopes of down-regulating the intensity of their negative affect (Rimé, 2009) or solicit alternative perspectives from their romantic partners after a major life event to assist their efforts to reappraise the situation as less threatening (Horn & Maercker, 2016). In this sense, IER acts as a possible lens to understand a variety of conceptually related social-emotional phenomena, including social and emotional support (e.g., Burleson, 2008; Lorenzo, Barry, & Khalifian, 2018; Maisel & Gable, 2009; Zee, Bolger, &

Higgins, 2020; see also Marroquín, 2011), venting and social sharing (e.g., Dixon-Gordon et al., 2018; Nils & Rimé, 2012; Parlamis, 2012; Rimé, 2009), reassurance-seeking (e.g., Abe & Nakashima, 2020; Dixon-Gordon et al., 2018), and dyadic coping (e.g., Bodenmann, 1995; O'Brien et al., 2009)—as efforts to engage in IER (see also Williams et al., 2018).

In emphasizing the complexity of the regulatory challenge involved in IER, Emily Butler and colleagues (2014) point out that IER frequently entails simultaneously "managing one's own emotions, a partner's emotions, and the emotional tone of the relationship" (p. 907). For example, a husband might suppress his own feelings of anxiety in service of helping his spouse manage their feelings around an upcoming anxiety-inducing situation. In the short-term, this exchange might yield pro-hedonic affective consequences for the spouse (e.g., feeling less worried), but contra-hedonic affective consequences for the husband (e.g., feeling more frustrated). If this pattern were to repeat many times, we might imagine that it could lead to increased conflict in the relationship in the long run—perhaps with the husband feeling that he is receiving inadequate support.

IER, then, is regarded as a prototypically reciprocal-transactional phenomenon insofar as it entails a series of motivated exchanges among multiple individuals that unfold over time (see Bodenmann, 1995; Butler, 2015). Not only are the outcomes of such exchanges likely to be contingent on extant qualities of the relationship and on the traits, states, and actions of each person involved, but also each individual's outcome is likely to be contingent on the other's outcome. For example, if the receiver of an IER interaction views the exchange as having been relatively successful, they might be more likely to express gratitude for their interaction partner, which might, in turn, tend to lead to more positive outcomes for the provider and for the relationship. Put another way, the consequences of providing and receiving IER are likely to be mutually interdependent, such that the behaviors enacted and the outcomes experienced by the provider are contingent on the behaviors and outcomes of the receiver and vice versa (see Butler, 2011; Butler, 2015; Nils & Rimé, 2012).

It is within this framework of mutual dependency that Katherine Dixon-Gordon and colleagues (2015) have discussed the challenge posed by IER in terms of a temporally unfolding, reciprocal process of encoding and decoding, in which both members of the dyad are simultaneously called upon to effectively communicate their own thoughts, feelings, and goals and to accurately interpret and respond to their partner's communications. This model resonates with findings from a preliminary fMRI investigation of IER that found that the neural correlates of providing IER (compared to the correlates of engaging in *intra*personal emotion regulation) encompassed brain regions that have been associated with social cognition and empathy (Hallam et al., 2014). Breakdowns at any stage of encoding and decoding would be expected to lead to poorer outcomes, such as selecting a strategy that is inappropriate for the situation (e.g., Wenzel, Rowland, Weber, & Kubiak, 2020; Shu, Bolger, & Ochsner, 2020) or misconstruing the others' motives (Niven, Troth, & Holman, 2019).

Collectively, these theoretical perspectives on IER suggest that IER may be a useful conceptual lens onto a plethora of deeply meaningful socio-emotional processes, but also one that may be challenging to capture and model adequately given the underlying dyadic and dynamic complexities. In the next section, I briefly review the extant IER literature with an emphasis on the utility of the construct vis-à-vis associations with functional outcomes.

Validity of IER: Mood Repair, Social Connection, and Psychological Well-being

Findings of multiple recent studies suggest that IER is routinely employed and can be an effective modulator of emotional experience throughout the lifespan. In a now classic daily diary

study, for example, college students endorsed regularly engaging in social sharing for the purpose of regulating both positive and negative emotions in their daily lives, and they rated this strategy as having as much of or even more of an impact on their affective states compared to various intrapersonal emotion regulation strategies (Heiy & Cheavens, 2014). In a similar vein, a pair of ecological momentary assessment (EMA) studies found that romantic partners frequently used physical touch and positive humor to regulate one another's emotions in daily life, and use of these behaviors was associated with increased momentary positive affect (Debrot, Schoebi, Perrez, & Horn, 2013; Horn, Samson, Debrot, & Perrez, 2018). Laboratory studies have yielded converging results. In one study, for example, romantic partners were able to effectively assist their partners efforts to down-regulate negative emotions in the context of a laboratory emotion regulation decision-making paradigm (Levy-Gigi & Shamay-Tsoory, 2017; see also Sahi, Ninova, & Silvers, 2021). In another, mother-child co-reappraisal significantly reduced the intensity of children's anger and sadness displays in response to a laboratory manipulation of disappointment (Morris, Silk, Morris, Steinberg, Aucoin, & Keyes, 2011). In other words, both naturalistic and experimental investigations support the notion that IER can exert a powerful influence on affective states across a swath of relational and developmental contexts.

Considering that IER is inherently social, it is unsurprising that multiple studies have investigated relationships between IER and interpersonal functioning. In one of the most thorough investigations of IER and interpersonal functioning, individuals who reported seeking IER more frequently also reported feeling more socially connected, were more likely to choose to view emotionally evocative images with another person (vs. than by themselves), and developed a larger network of more supportive relationships when followed through the first year of college (Williams et al., 2018). In a similar vein, the tendency to engage in more co-rumination has been found to be associated cross-sectionally with high quality, close friendships and to predict longitudinally the development of closer friendships among children and adolescents (Rose, 2003; Rose, Carlson, & Waller, 2007). In two studies of daily life, the use of physical touch and positive humor by romantic partners in service of IER was associated with increased feelings of psychological intimacy in the relationship (Debrot, Schoebi, Perrez, & Horn, 2013; Horn, Samson, Debrot, & Perrez, 2018). There is some evidence that the relational consequences of receiving IER may depend on the type of IER that is provided; in one study, for example, receiving socio-affective support (e.g., comfort and validation) in response to social sharing of emotions was significantly associated with increased interpersonal closeness with the listener, whereas receiving cognitive support was not (Pauw et al., 2017). Taken together, multiple facets of IER appear to be tied to social connectedness and close social bonds.

IER may also have important ramifications for psychological health and well-being (see Hofmann, 2014; Lopez-Perez, Ambrona, & Gummerum, 2016; Marroquin, 2011). For example, reporting a greater number of specialized "emotionships" (i.e., relationships that fulfill emotion-specific emotion regulation needs, such as having one friend that you tend to turn to when you are feeling sad and another for when you are feeling angry) was associated with higher life satisfaction, even after adjusting for the number of close relationships (Cheung, Gardner, & Anderson, 2015). Among romantic couples who had experienced a major life stressor in the past year, tendencies to engage in co-brooding were related to greater preoccupation with the stressor, greater ongoing functional impairment related to the stressor, and more severe symptoms of depression, whereas co-reappraisal tendencies were associated with lower scores on these measures of adjustment difficulties (Horn & Maercker, 2016). Similar patterns of associations have been observed in studies of close friendships (Christensen, van Dyk, Nelson, & Vasey,

2020; Spendelow, Simonds, & Avery, 2016). Finally, in a laboratory study of mother-daughter conflict conversations, maternal problem-solving in the absence of maternal validation was associated with more severe borderline personality disorder symptoms (Dixon-Gordon et al., 2016). Of import, then, as with intrapersonal emotion regulation, some types of IER patterns appear to be tied to higher levels of psychological health and well-being across individuals, whereas others appear to be tied to lower levels.

Findings of several longitudinal studies have likewise suggested that IER is tied to mental health. Across two longitudinal studies, adolescents' co-rumination tendencies positively predicted internalizing symptoms and depressive episodes (Rose, Carlson, & Waller, 2007; Stone, Hankin, Gibb, & Abela, 2011). In a similar vein, inclinations to engage in avoidance and reassurance-seeking in response to hypothetical interpersonal emotional scenarios predicted greater physical aggression and self-injury, respectively, over the course of a two-week daily diary study (Dixon-Gordon et al., 2018). Collectively, these findings suggest that IER patterns are tied to psychological functioning.

Beyond receiving IER, the act of providing IER has likewise been shown to have important implications for affective and social functioning (e.g., Niven, Holman, & Totterdell, 2012; Niven, Macdonald, & Holman, 2012; Martínez-Íñigo, Poerio, & Totterdell, 2013; Niven et al., 2015; Doré et al., 2017). For example, the provision of IER has been tied to popularity (Niven et al., 2015), psychological well-being (Niven, Macdonald, & Holman, 2012), and the strategies that people use to regulate their own emotions (Doré et al., 2017).

Preliminary IER Work by Our Team

Within the broader context of this emerging IER literature, my colleagues and I have conducted research with the aim of understanding the structure and outcomes of IER interactions. To these ends, we have conducted a series of cross-sectional self-report, autobiographical recall (e.g., recall a recent time when another person tried to change how you were feeling), and daily diary studies in undergraduate, community, and clinical samples (Swerdlow, Berk, & Johnson, *in prep*; Swerdlow & Johnson, 2020; Swerdlow & Johnson, 2021; Swerdlow & Johnson, *in prep*; Swerdlow, Sandel & Johnson, *in press*).

Broadly speaking, our findings have aligned with and extended prior IER research. Our data suggest that IER is routinely employed in the daily lives of college students and that college students tend to report that receiving IER is at least as helpful in improving their affective states (i.e., making them feel more positive and less negative) as their intrapersonal emotion regulation efforts. In addition, we find that instances of IER that are experienced as affect-improving by the receiver also tend to be rated by the receiver as increasing their ability to cope, their sense of control over their emotions, their positive feelings about both themselves, their sense of connectedness with their IER provider, and their likelihood of seeking IER in the future. That is, retrospective ratings of these interaction outcomes tend to be highly correlated (Swerdlow & Johnson, 2020). This finding suggests not only that people perceive IER interactions as conveying multiple benefits above and beyond hedonic repair, but also that people tend to regard individual IER interactions as relatively helpful or unhelpful at the level of a gestalt (opposed to discriminating between specific outcomes). We refer to this gestalt as the "perceived benefits" of the interaction—as distinct from "effectiveness," as the latter might be taken to connote the "objective" extent of affective change or adaptation.

People's inclinations to seek IER and their perceptions of the benefits of IER interactions appear to be tied to their *intra*personal emotion and emotion regulation tendencies. For example, data collected by our team and by others (e.g., Williams et al., 2018) suggest that people who are

more liable to use expressive suppression to regulate their emotions also tend to be less apt to want or seek IER (Swerdlow & Johnson, 2021). However, these effect sizes are generally modest, such that ratings of intrapersonal emotion tendencies and of experiences of seeking and receiving IER are substantially distinct (Swerdlow & Johnson, 2020; Swerdlow & Johnson, 2021). In a similar vein, we have found that people with clinically relevant difficulties with mood and emotion regulation often have difficulty finding helpful IER. Compared to a sample of college students, for example, a sample of adults seeking treatment for emotionally impulsive aggression reported seeking IER less frequently, having more difficulty finding willing IER providers when they did seek it, encountering more hostility and less responsiveness from IER providers, and finding IER to be less helpful (Swerdlow, Berk, & Johnson, *in prep*). Parallel findings were obtained for ratings of global experiences of IER and of an autobiographically recalled instance of receiving IER. Within a sample of college students, current symptoms of anhedonic depression were associated with similar difficulties finding helpful IER (Swerdlow & Johnson, 2016). In absolute terms, though, all of the samples that we have studied have rated IER interactions as at least somewhat beneficial, on average. In summary, whereas some people may tend to experience greater challenges related to receiving IER than others, it appears that IER interactions can often still be helpful for those who struggle to regulate their own emotions.

We have also considered specific facets of IER interactions and the outcomes tied to those facets. More specifically, we developed a self-report measure, the Interpersonal Regulation Interaction Scale (IRIS), that captures receivers' perceptions of four factor-analytically derived dimensions of IER interactions: provider responsiveness (i.e., understanding, validating, and caring about the receiver), provider hostility (i.e., responding harshly, critically, or dismissively), provider cognitive support (i.e., providing an alternative perspective or problem-solving), and provider physical presence (i.e., communicating availability and non-verbal expressions of warmth and affiliation; Swerdlow & Johnson, 2020). As above, these perceptions were modestly associated with, but substantially distinct from, intrapersonal emotion regulation tendencies, as well as an array of other individual differences (e.g., personality). Moreover, the degree to which an IER interaction was perceived as beneficial by the receiver was robustly tied to the degree to which the provider was perceived as responsive by the receiver and, to a lesser extent, to perceptions of hostility and cognitive support (Swerdlow & Johnson, 2020). In one undergraduate daily diary study, though, daily psychological well-being and daily loneliness were most strongly tied to the extent to which providers were perceived as having conveyed more cognitive support and more hostility, respectively, suggesting that there may be some specificity to the links between various facets of the interaction and specific psychological outcomes after all (Swerdlow & Johnson, *in prep*).

Alongside the possibility of increased loneliness, our research also points to at least one other potential pitfall of seeking and receiving IER: namely, that even instances of IER that are recalled as having been helpful overall can simultaneously be recalled as having been shame-inducing. Across six samples using a variety of self-report methods (total $n = 1868$), between 15% and 33% of reported instances of receiving IER were associated with moderate or greater endorsement of shame as a result of receiving IER (total $k = 2515$), and these shame ratings were only moderately correlated with ratings of the helpfulness of the IER interaction (Swerdlow, Sandel, & Johnson, *in press*). In an internal meta-analysis of these samples, we found that ratings of shame in this context were robustly correlated with individual differences in nonacceptance of negative emotions, use of expressive suppression, shame-proneness, and negative affectivity; with lower situational desire for IER; with having a less close relationship with the IER provider;

and with perceptions of the IER providers as having been more hostile and less responsive (Swerdlow, Sandel, & Johnson, *in press*).

Whereas these data suggest the *possibility* that specific provider behaviors can be linked to specific socioemotional outcomes, a crucial limitation of these data is that they reflect receivers' perceptions of providers' behavior and not the behaviors themselves. On the one hand, receivers' perceptions may be more proximal predictors of their psychological outcomes than providers' behaviors *per se*. For example, our interpretation of another person's behavior as responsive may exert more direct influence over whether we feel more closely connected to them than the number of "objectively" responsive utterances (cf. Reis & Gable, 2015). On the other hand, these perceptions are inherently confounded insofar as they represent an uncertain amalgam of enacted behavior and appraisal, with the latter opening the door to an array of attentional and interpretational biases. As one example of this challenge, in our research, we have consistently found that higher symptoms of depression cross-sectionally correlate with and longitudinally predict perceptions of IER providers as less responsive and more hostile. From these data alone, it is not possible to determine whether those associations are driven primarily by negative cognitive biases associated with depression or by depressed individuals actually eliciting less responsive and more hostile responses, although both of these interpretations would find long histories of support in the depression literature (e.g., Gotlib, 1983; Hokanson et al., 1989). This observation of the behavior-appraisal confound constitutes one of the central rationales for the present investigation, and so I devote the next section to a more detailed articulation of the problem.

Behaviors vs. Appraisals

Within the close relationships literature, there are rich empirical traditions of modeling the contributions of couples' behaviors and appraisals to interpersonal and psychological outcomes. Important questions persist, though, as to whether behaviors or appraisals carry the most predictive weight with respect to any given outcome and how exactly behaviors and appraisals are related to one another.

Multiple longitudinal studies published over the past three decades have highlighted the predictive power of the enacted behaviors that couples exhibit in the laboratory during problem-solving and conflict conversations across a variety of intrapersonal and interpersonal outcomes, including relationship satisfaction and conflict, divorce, physical health, morbidity, and longevity (e.g., Bloch, Haase, & Levenson, 2014; Gottman, Coan, Carrere, & Swanson, 1998; Gottman & Levenson, 2000; Haase et al., 2017; Julien, Markman, & Lindahl, 1989; Kiecolt-Glaser, Bane, Glaser, & Malarkey, 2003; Otero et al., 2020; Pietromonaco, Overall, & Powers, 2021; Wells et al., 2022). Famously, John Gottman's identification of the "four horsemen of the apocalypse"—four behaviors that were found to predict marital dissatisfaction and divorce across couples (i.e., criticism, defensiveness, contempt, and stonewalling)—became a foundational tenet of The Gottman Method, a widely used couples therapy (Gottman & Gottman, 2008). Of import, the available evidence suggests that these behavioral patterns can remain relatively consistent over time within long-term romantic relationships (Gottman & Levenson, 1999). This may explain in part how a single brief snapshot of a couples' communication style—even one obtained in a laboratory—can predict long-term outcomes. Notwithstanding that these communicative behaviors may act, in part, as proxies for or symptoms of a range of intrapersonal and relational vulnerabilities and difficulties, this body of research powerfully highlights both the scientific and practical utility of adopting a behavioral lens.

An enduring critique of these behavioral approaches, though, is that they do not account for appraisals that might alter the interpretation of behaviors (see Bradbury & Fincham, 1990). For example, the tendency to extend generous appraisals within marital relationships (e.g., tending to make benign as opposed to nonbenign attributions regarding the causes of partners' undesired behaviors) is associated with and predicts marital quality (e.g., Bradbury & Fincham, 1990; Bradbury & Fincham, 1992; Bradbury, Beach, Fincham, & Nelson, 1996). Indeed, it has been suggested that positive illusions may be an integral component of happy relationships (Murray, Holmes, & Griffin, 1996a; Murray, Holmes, & Griffin, 1996b; Murray & Holmes, 1997; see also Fletcher & Kerr, 2010). Partners' perceptions of responsiveness—defined as the degree to which partners perceive one another as understanding, caring, and validating—has been proposed to be a unifying proximal predictor of various relationship outcomes (Reis & Gable, 2015).

Although these more cognitively oriented theories do not discount the role of behavior altogether, they do potentially argue against robust behavior-outcome relationships insofar as the consequences of behaviors are thought to be subjected to the moderating influence of motivated and potentially relationally idiosyncratic appraisals. In fact, several recent studies have suggested that inter-rater agreement regarding the supportiveness of other people (Lakey, 2010) and of supportive statements (Tanner et al., 2018) is strikingly low (see also Lakey & Orehek, 2011). To the extent that this is the case, it might be important to develop interventions aimed at helping couples shift their cognitions and see their partners' actions in a different light rather than focusing solely on altering behaviors or communication patterns (e.g., Margolin & Weiss, 1978; Dattilio & Padesky, 1990).

Behaviors and appraisals are generally linked to one another in at least a loosely coherent fashion (cf. Bradbury & Fincham, 1990; Reis & Gable, 2015). As one example, research on responsiveness has found that partners' perceptions of responsiveness are correlated with independent coders' ratings of enacted responsiveness (Reis & Gable, 2015), suggesting that while motivated cognition may play a role, relationship appraisals are nevertheless grounded in concrete behavior to some degree. In turn, partners' appraisals may give rise to behaviors, as in the case of an ungenerous appraisal of a partner's behavior leading to more enacted criticism (Sanford, 2006). Ultimately, then, dyadic studies that measure both behavior and appraisal offer a particularly powerful lens into relationship dynamics. One compelling example is a dyadic laboratory study of social support, which found that the benefits of social support were maximized by a combination of providers' skillful delivery (as rated by independent coders) and receivers being unaware of having received support (i.e., low support visibility as rated by the receivers; Howland & Simpson, 2010). Accordingly, one of the central goals of the current investigation was to evaluate an ecologically valid paradigm for the simultaneous assessment of interpersonal regulation behavior and subjective experiences of interpersonal regulation.

The Present Study

The forgoing discussion highlights the utility of adopting a dyadic perspective and of simultaneously measuring behavior and subjective experience to provide richer accounts of IER. I am also committed to studying IER within an ecologically valid framework that reflects the actual contingencies and correlations that structure individuals' real social environments (e.g., the contexts in which, the individuals from whom, and the strategies whereby people actually receive and provide IER, and the meaningful interrelations among these features). Therefore, in the present study, my team and I tracked daily IER interactions in a community sample of committed, cohabiting adult romantic couples using both EMA to capture subjective experience

(Shiffman, Stone, & Hufford, 2008) and ambulatory acoustic recording to capture enacted social behavior (Mehl, 2017).

I made the decision to focus specifically on adult romantic couples not because they are the only stable dyad in which IER routinely occurs, but rather because I view romantic couples as a model dyad in which to study IER. There are several reasons to regard cohabiting romantic couples as a model dyad. First, I anticipated a relatively high base rate of IER due to the combination of high physical proximity and psychological intimacy. This was important, among other reasons, to have a sufficiently high number of IER interactions for statistical analyses. Second, committed romantic couples are likely to have developed habitual patterns of IER (cf. Gottman & Levenson, 1999), such that sampling even a subset of their IER interactions ought to provide a relatively powerful window into their relationship dynamics. Third, I anticipated that the consequences of IER for psychological health would be magnified by the primacy of the romantic relationship context (cf. Shaver & Mikulincer, 2007). That is, a romantic partner's routine failure to provide responsive IER might have more significant and wide-ranging consequences than that of an acquaintance or even a close friend. Finally, in contrast with parent-child dyads, I anticipated more mutuality of IER, which aligned with the goal of examining the consequences of both receiving and providing IER.

A major question from the outset was how best to capture experiences of IER within an ecologically valid framework. As previously noted, our prior research on IER has relied on self-report, and I have already noted one limitation of this approach, which is that behaviors and appraisals cannot be disentangled. A closely related limitation is that we have been limited exclusively to studying interactions that participants (retrospectively) classified as IER. Drawing on the social support literature and especially on the distinction between visible and invisible social support (see Zee & Bolger, 2019), one major question that has not yet been addressed by IER researchers to our knowledge is whether receivers' awareness of IER interactions as such has an impact on the outcomes.

Therefore, in addition to completing multiple brief self-report assessments on their smartphones each day, participants wore digital audio recorders, which permitted my team and I to "listen in" on their daily interactions with their partners (see Mehl, 2017 for a discussion of ambulatory acoustic sampling to measure real-world social behavior). Ambulatory acoustic recording has been used previously to study dyadic coping conversations in couples coping with breast cancer (Robbins et al., 2014; Robbins, Wright, López, & Weihs, 2019), as well as other social and communicative behaviors (e.g., Herbison et al., 2021; Minor et al., 2018; Robbins et al., 2011; Slatcher & Robles, 2012). Prior validations of ambulatory acoustic recording suggest that most people habituate rapidly to the technology and perceive it as reasonably unobtrusive (e.g., Mehl & Holleran, 2007; Manson & Robbins, 2017), and this method has been successfully used in a wide variety of contexts, including both non-clinical and clinical populations and across the lifespan (e.g., Baddeley, Pennebaker, & Beevers, 2012; Demirary, Luo, & Grilli, 2020; Minor et al., 2018; Slatcher & Robles, 2012; Tomko et al., 2014). We are not aware, though, of any published studies that have used ambulatory acoustic recording expressly to study IER, and ambulatory acoustic recording has not typically been combined with momentary self-report. In the current study, acoustic recordings were rated by trained undergraduate content coders using a manual that I developed for this purpose (see Appendix A). Ultimately, I view the self-report and behavioral acoustic data as essentially complementary: providing us with a more objective source of information about the interactions in questions (i.e., allowing us to parse the enacted behavioral content), while simultaneously retaining access to participants' perceptions of

themselves, their experiences, and their relationship. In sum, to my knowledge, this study is one of only a handful of dyadic IER studies to use ecological sampling methods of any kind, and the first to use ambulatory acoustic recording.

Central Aims

The ultimate aim of this research program is to understand how IER interactions transact with socioemotional outcomes. At the broadest level, then, I aim to investigate the dyadic, interdependent relationships among receiving and providing IER, momentary experienced affect, psychological well-being, and relationship quality. This broader aim rests on establishing that we can feasibly collect data using a combination of EMA and ambulatory acoustic recording and reliably identify and code the content of IER interactions in the acoustic data. Therefore, the central aims of the current project were to evaluate the feasibility, reliability, and preliminary validity of the data collection and coding apparatus that we developed. Each of those aims is detailed below.

Aim 1 (Feasibility). Evaluate technical and nontechnical barriers of using the EAR for Android app as a tool for collecting ecologically valid IER data in a dyadic context. Although the use of acoustic recording as an ecological data collection tool is not new (e.g., Hart, Risley, & Kirby, 1997), to the best of my knowledge, this study was the first to use ambulatory acoustic recording as a tool for studying IER and one of only a small handful that have used ambulatory acoustic recording in a dyadic context or alongside self-report EMA. In addition, the EAR for Android smartphone app, which is the latest implementation of the Electronically Activated Recorder (EAR; see Mehl, 2017), is actively under development. Specifically, I considered: 1) technical and procedural barriers, including storage capacity, battery life, system errors, legal and ethical considerations, and possibilities for automated transcription; 2) acceptability and obtrusiveness of the EAR, including the number of participants who drop out of the study; 3) the number of recordings produced, as well as variability in the number of recordings produced across participants; and 4) the base rate of audible IER interactions occurring within the couples.

This last consideration is particularly important given that there may be a number of reasons that we might "miss" or otherwise exclude genuine instances of interpersonal regulation regardless of the particular implementation of acoustic recording. These reasons include: 1) nonverbal interpersonal regulation, including interpersonal regulation that occurs via a text-based medium; 2) interpersonal regulation that takes place outside of the hours during which we are recording or when participants are otherwise not wearing the EAR, potentially because they are uncomfortable discussing emotionally sensitive topics while wearing the device; or 3) instances of interpersonal regulation that are audibly accessible, but for which we cannot establish with adequate certainty that the dyad comprises one of our couples (e.g., if one of the participants' receives interpersonal regulation from someone other than their romantic partner). In the current study, we applied a periodic sampling strategy to the acoustic data in consideration of the total number of hours of audio that we planned to collect, which would also be expected to contribute to "missing" IER interactions.

Aim 2 (Reliability). Ascertain the reliability and structure of our coding scheme. As we were applying a novel coding scheme to the EAR recordings, an essential aim was to examine the reliability and structure of those codes. First, I planned to establish the degree to which coders could reliably identify the phenomenon of interest (i.e., interpersonal regulation present vs. interpersonal regulation absent) using the operational definition that I created (see Appendix A for details). For those interactions that could not be reliably identified (i.e., one coder indicated that interpersonal regulation was present, but a second coder indicated that interpersonal

regulation was absent), we used an inductive approach to identifying potential conceptual/content gaps in our operational definition to inform future studies, including identifying the most common challenges that coders encountered. For those interactions that were reliably identified as interpersonal regulation, I planned to calculate reliabilities for the content codes and identify codes that did not achieve acceptable reliability. Finally, for all of the codes that achieved acceptable reliability, I planned to examine univariate distributions and bivariate correlations and to submit those codes to an exploratory factor analysis to begin to understand the statistical properties and structure of this coding system.

Aim 3 (Validity). Conduct an initial assessment of the validity of our coding scheme. To this end, I planned to conduct several specific tests of validity and utility:

3a. To examine correspondence between participants' self-report EMA and coder ratings of their acoustic data. More specifically, I examined correspondence not only in the incidence and timing of IER interactions, but also in ratings of the content of those interactions. I anticipated that these correlations would be significant, but modest. One goal of these analyses was to calculate the frequency with which our team was capturing IER interactions that were not visible to the participants as such (i.e., when the participants reported *not* having engaged in IER, but we observed IER).

3b. To examine predictors and correlates of our ratings of enacted provider behavior. First, I examined whether these ratings could be predicted from key baseline measures. Then, I examined whether the frequency of IER interactions tracked with daily self-reported stress and relationship quality. More specifically, we hypothesized *a* priori that IER would occur most frequently in couples who were experiencing heightened stress, but who also perceived one another as highly responsive.

Method

All procedures were approved by the university Institutional Review Board before data collection. We applied for and received a Certificate of Confidentiality from the National Institutes of Health prior to study launch as an additional safeguard for participant privacy.

Participants

Couples were recruited primarily from the greater San Francisco Bay Area. We recruited couples online and using flyers, including outreach to local organizations that serve couples and families. Efforts were made throughout recruitment to ensure that the sample would be representative of the Bay Area in terms of ethnicity and socioeconomic status (e.g., by posting flyers at organizations that primarily serve minoritized and marginalized communities). All participants were required to: 1) be 18 years of age or older, 2) be in a committed, cohabiting romantic relationship, 3) have regular access to a smartphone to complete EMA surveys, 4) indicate their willingness to be audio recorded as part of the study, 5) share two or more "weekend days" in common with their partner, 6) predominantly speak English when communicating with their partner at home, and 7) indicate their willingness to travel to the university campus where the study was administered. For either member of a couple to participate, both members had to meet these criteria and consent to participate in the study. No restrictions were placed on participants' sexual orientations (i.e., participating couples did not have to be opposite-sex or opposite-gender), nor were participants required to be legally married.

Between October 2018 and May 2020, there were 772 completed responses to a brief online screener for the study. Of these, 375 (48.6%) were deemed ineligible to participate. By far, the most common reason for ineligibility was respondents indicating that they were not currently cohabiting with their romantic partner (320 responses, or 85.3% of ineligible

responses). Less commonly endorsed reasons for exclusion included: not sharing two or more "weekend days" in common with their partner (45 respondents [12%]), not being in a committed romantic relationship (40 [10.7%]), being unwilling or unable to travel (16 [4.3%]), not primarily speaking English at home with their partner (16 [4.3%]), their partner not being interested in participating (15 [4%]), missing an attention check that was included in the screener (13 [3.5%]), being unwilling to be audio recorded as part of the study (5 [1.3%]), being below the age of 18 (4 [1.1%]), and not having regular access to smartphones (2 [0.5%]).

Ultimately, 63 couples attended an intake and consented to participate in the study. Of those who attended an intake and declined to consent, most stated concerns about the time commitment or payment amount, not the audio recordings. Of those who consented to participate, four couples did not complete any of the three major components of data collection (i.e., baseline questionnaires, audio recordings, EMA surveys). Among the remaining 59 couples, most couples completed all three major study components. That is, all but one couple (98%) completed baseline questionnaires, all but four couples (93%) submitted ≥ 1 audio recordings, and all but 6 couples (90%) submitted ≥ 1 EMA surveys.[1]

This participating sample (n = 118) was split approximately in half by gender (46.4% female, 50.9% male, 2.7% non-binary or genderqueer). Most couples (88.2%) were opposite-gender. Participants identified with the following ethnicities: 42.9% White or Caucasian, 28.6% Asian or Asian-American, 10.7 Hispanic or Latino/a/x, 9.8% multiple ethnicities, 6.3% Black or African American, 1.8% wrote in another ethnicity or declined to respond. Participants ranged from 18 years old to 70 years old (M = 29.30, SD = 11.05). The modal participant had completed a college degree (M = 15.6, SD = 3.24). A plurality of participants were employed full-time (37.5%); the remainder were full-time students (34.8%), employed part-time (17.9%), looking for work (8%), or stay-at-home parents (1.8%). The median couple had been together about two and a half years, but there was considerable variability in relationship duration (M = 63.50 months, SD = 95.43 months). About a third of participants (34.8%) were legally married or in domestic partnerships. Most participants (75.6%) did not have any children with their partner.

Participant Procedures

As detailed above, prospective participants completed an online prescreening questionnaire to determine their probable eligibility for the study. Those who were designated likely eligible were invited to schedule an in-person or remote[2] intake, during which they completed written informed consent procedures, a baseline questionnaire battery, two computerized behavioral tasks, and an orientation to the audio recorder and the smartphone app that we used to collect data.

Participants were asked to respond to EMA surveys and to wear the audio recorder for four consecutive weekends (defined as two days out of the week during which the couple

---

1 One likely contributor to participant nonadherence was the COVID-19 pandemic. Of the 63 couples who were consented into the study, 48 (76%) were consented and scheduled to complete the study before March 11, 2020—the date of the WHO pandemic declaration. Of the four couples that did not complete any of the three major components of the study, though, three of those couples (75%) were consented into the study or were scheduled to complete the study *after* March 11, 2020. Likewise, of the 6 couples who did not complete any EMA records, 3 (50%) were consented into the study or scheduled to complete the study after March 11, 2020.

[2] After the advent of the COVID-19 pandemic, intakes were conducted remotely via HIPAA-compliant Zoom and DocuSign.

typically did not spend most of the day at work and instead tended to spend more of their time together), for a total of eight days of data collection. This scheme was devised to obviate concerns that participants might have had about wearing the audio recorder at work and to increase the expected rate of couples' interactions.

On each day when EMA data were collected, participants received four identical surveys at pseudorandomly spaced intervals throughout the day (between the hours of 10:00 AM and 8:00 PM) via the RealLife Exp. Participants also completed a nightly diary at the end of each of those same days (i.e., 10:00 PM). The time window for the EMA surveys was designed to match that used for the acoustic recording (see below). It was not apparent, *a priori*, what the frequency of IER interactions in romantic might be, but five surveys per day is a commonly used sampling rate for this method (see Hall, Scherner, Kreidel, & Rubel, 2021 for review). The median response time for completed EMA surveys was under two minutes (88 seconds). On average, participants completed 77% of EMA surveys, which is typical for EMA research (e.g., Kaurin, King, & Wright, 2022).

The digital audio recorders were Android smartphones provided by the researchers with the EAR for Android app developed by Matthias Mehl and colleagues pre-installed and configured on them (Mehl, 2017). The audio recorders were programmed with a "blackout period" so that they only recorded between the hours of 10:00 AM and 8:00 PM local time but were otherwise programmed to record continuously throughout this interval in ten-minute segments. We followed standard procedures for instructing participants in the use of the EAR (see Robbins, Wright, Karan, & Baranski, 2021). Specifically, participants were asked to incorporate the audio recorders into their daily lives as naturally as possible, not to conceal the audio recorders from others, and to wear the audio recorders as much as possible during their waking hours. Participants were given a button to wear in public to inform bystanders of the possibility that they would be audio recorded to mitigate privacy-related legal and ethical concerns.[3] Participants were informed that they could take the audio recorder off at any time and that there were times when they should take the audio recorder off (e.g., if they were showering, if another person objected to being recorded). At the same time, the importance of and corresponding rationale for wearing the audio recorder as much as possible was communicated verbally and in written instructions. Per standard EAR protocol, all participants were given the opportunity to listen to their recordings before data coding and to indicate that they wanted one or more of their audio recordings not to be processed or coded. These requests were honored without question or exception. Participants received reminders on their smartphones via the RealLife Exp app in the morning (to turn on and wear the EAR devices) and in the evening (to turn off and charge the EAR devices) on their recording days.

After they completed the requisite eight days of data collection, participants were emailed a link to an end-of-study questionnaire battery that included repetitions of several of the baseline questionnaires, as well as a questionnaire about their experience of wearing the EAR. After the onset of the COVID-19 pandemic, participants who had previously completed the study were

[3] California is an all-party (AKA two-party) consent state, meaning that California law requires the consent of everybody involved in a conversation before the conversation can be legally recorded. The button establishes that people other than the participants who may come into contact with the participants when they are wearing the audio recorder can reasonably expect that their communication may be overheard or recorded (see California Penal Code § 632(a)-(c)). See also, Manson & Robbins, 2017.

invited to complete an optional, unplanned follow-up questionnaire battery to assess their individual and relational functioning amid the pandemic; however, data from this follow-up was not included in the present analyses, and so will not be discussed further.

Participants were paid for their time and effort up to a maximum of $72 per participant ($144 per couple). These payments included specific incentives for completing the EMA surveys. Participants were also reimbursed for travel expenses (e.g., parking, public transportation). If the EAR devices had to be returned by mail, participants were provided with a pre-stamped envelope or reimbursed for the cost of shipping.

Measures

A complete inventory of study measures was pre-registered on the OSF (https://osf.io/974cn/).

Baseline

Participants completed two laboratory tasks and a battery of questionnaires at baseline to measure key variables, including intra- and inter-personal emotion regulation, relationship quality, and psychological well-being, among others.

Interpersonal emotion regulation. At baseline, participants completed a questionnaire designed to assess IER tendencies and preferences. To emphasize the goal-directed aspect of IER, IER was defined as times "when someone did or said something to try to help you manage your emotions or feel better," and participants were instructed to "think about times when you wanted to feel more or less positive, more or less negative, or more or less calm, and someone else tried to help you." To assess stable individual differences, participants were instructed to "think about what you generally do and feel." Participants were first asked a series of five questions about the frequency (e.g., "How often do you actively seek it out"; 1 = never or almost never; 5 = very often) and helpfulness (e.g., "how often does it help you to change how you are feeling?"; 1 = never or almost never; 5 very often or almost always) of IER.

Participants then completed a modified version of the Interpersonal Regulation Interaction Scale (IRIS; Swerdlow & Johnson, 2020). The original IRIS asks respondents to rate their perceptions of a single IER interaction along four factor-analytically derived dimensions: provider responsiveness, provider hostility, provider cognitive support, and provider physical presence. In the current study, participants rated identical items, but with a modified item stem and response options: "In general, when I'm receiving [IER], I prefer it when the other person…" (e.g., "gives me advice", "lets me vent my emotions"). Items were rated on a five-point scale (1 = doesn't do this at all; 5 = does a lot of this). Items were averaged to form the four subscales. Internal consistencies for the four subscales were good for Cognitive Support (Cronbach's $\alpha = .87$), Responsiveness (Cronbach's $\alpha = .85$), and Physical Presence (Cronbach's $\alpha = .80$) and acceptable for Hostility (Cronbach's $\alpha = .76$).

Relationship quality. Participants completed the Couples Satisfaction Index-16 (CSI-16; Funk & Rogge, 2007) at baseline. The CSI-16 is a well-validated, unidimensional self-report measure that is specifically designed to measure relationship satisfaction (opposed to, e.g., measuring relationship conflict or communication). The CSI-16 was constructed using item-response theory and principal components analysis to identify the fewest items that provided the most precise information about relationship satisfaction from an initial pool of 176 items drawn from common measures of relationship satisfaction. Exemplar items include: "Please indicate the degree of happiness, all things considered, of your relationship" and "I have a warm and comfortable relationship with my partner." Prior research indicates that the CSI-16 possesses strong construct and convergent validity (Funk & Rogge, 2007). Items were summed to create a

single composite relationship satisfaction score, which showed excellent internal consistency in this sample (Cronbach's $\alpha = .96$).

Participants also completed several measures designed to tap into more specific relationship appraisals; those being: the Romantic Love Scale (Rubin, 1970), the Perceived Partner Responsiveness Scale (Reis et al., 2018), the Evaluation of Dyadic Coping subscale of the Dyadic Coping Inventory (Bodenmann, 2008).

Romantic love. The Romantic Love Scale (Rubin, 1970) is a 13-item measure designed to assess the respondent's affiliative orientation toward their romantic partner. The items were written with the aim of distinguishing romantic love from romantic liking (i.e., romantic attraction), where the former was theorized to comprise three components: affiliative and dependent need (attachment), predisposition to help (caring), and orientation to exclusiveness and absorption (intimacy). A factor analysis of several love scales concluded that the Romantic Love Scale taps into the facet of love often referred to as companionate love (opposed to passionate love; Fehr, 1994). Exemplar items include: "I feel that I can confide in my partner about virtually anything"; "I would do almost anything for my partner"; and "It would be hard for me to get along without my partner." All items were rated on a Likert scale (1 = not at all true/disagree completely; 9 = definitely true/agree completely). Items were summed to create a composite score, which showed good internal consistency (Cronbach's $\alpha = .82$).

Perceived partner responsiveness. Participants' perceptions of their partners' responsiveness were measured using the Perceived Partner Responsiveness Scale (Reis et al., 2018). The Perceived Partner Responsiveness Scale is an 18-item self-report measure designed to capture the extent to which people feel cared about, understood, and validated by their romantic partner. As previously discussed, perceptions of partner responsiveness have been proposed to be a core constituent of perceptions of relationship quality and effective relational communication (Reis & Gable, 2015). Perceived responsiveness has previously been found to ameliorate negative effects of visible received support (Maisel & Gable, 2009) and facilitate emotional expression in romantic relationships (Ruan et al., 2020). Exemplar items include: "My partner usually understands me" and "My partner usually esteems me, shortcomings and all." Participants responded to these items on a Likert scale (1 = not at all true; 9 = completely true). Consistent with the most common usage of the scale and with prior factor analytic evidence of unidimensionality (Reis et al., 2018), a total responsiveness score was computed by summing all 18 items. This total score showed excellent internal consistency (Cronbach's $\alpha = .98$).

Evaluation of dyadic coping. Participants indicated their satisfaction with how they and their partner cope with stress together using a modified version of the Evaluation of Dyadic Coping subscale of the Dyadic Coping Inventory (Bodenmann, 2007). The 37-item Dyadic Coping Inventory is the most commonly used instrument for measuring perceptions of dyadic coping with stress in romantic relationships. The full Dyadic Coping Inventory contains 10 subscales; in the current study, though, only the items from the Evaluation of Dyadic Coping subscale were administered. The original Evaluation of Dyadic Coping subscale consists of two items: "I am satisfied with the support I receive from my partner and the way we deal with stress together" and "I am satisfied with the support I receive from my partner and I find as a couple, the way we deal with stress together is effective." In the current study, though, these compound items were dismantled so that the modified scale consisted instead of three items: "I am satisfied with the way that my partner and I deal with stress together"; "I find that, as a couple, the way we deal with stress together is effective"; and "I am satisfied with the support I receive from my partner." Participants rated these items using a frequency scale (1 = very rarely; 5 = very often).

Participants' responses were summed to create a total score, which showed excellent internal consistency (Cronbach's $\alpha = .91$).

Intimate partner violence. A subset of participants completed a measure of experiencing intimate partner violence in their current relationship. We included 29 items adapted from the National Intimate Partner and Sexual Violence Survey (Kresnow, Smith, Basile, & Chen, 2022) that covered specific behaviors related to psychological/emotional aggression (e.g., "Told you that no one else would want you?"), coercive control (e.g., "Tried to keep you from seeing or talking to your family or friends"?), and physical aggression (e.g., "Hit you with a fist or something hard?"). Respondents were asked to indicate how often in the past twelve months their partner had engaged in each behavior on a frequency scale (0 = not in the past twelve months; 5 = daily or almost daily).

Only one member of each couple was asked to complete the questionnaire to enhance participant privacy and safety. Participants who were designated to receive the intimate partner violence questionnaire were informed that their partners would *not* be receiving the same questionnaire. Before receiving the questionnaire, participants were informed that the next questionnaire would entail questions about intimate partner violence and were asked to indicate whether they felt safe to answer questions about intimate partner violence. Only participants who responded to this question in the affirmative proceeded to the questionnaire.[4] As women are disproportionately affected by at least some forms of intimate partner violence (e.g., Caldwell, Swan, & Woodbrown, 2012; Jonas et al., 2014), in the case of male-female couples, only the female partner was designated to receive the intimate partner violence questionnaire. In the case of same-gender and genderqueer couples, one partner was randomly designated to receive the intimate partner violence questionnaire using a random number generator. All participants who were designated to receive the intimate partner violence questionnaire received a referral to the National Domestic Violence Hotline (https://www.thehotline.org/) regardless of their responses. These procedures were adopted consistent with ethical guidelines for the assessment of intimate partner violence in research settings (see Bender, 2017).

Recent internalizing symptoms, psychological well-being, and perceived stress. Participants completed selected subscales of the Inventory of Depression and Anxiety Symptoms-II (IDAS-II; Watson et al., 2012) to assess their mental health. The full IDAS-II contains 99 items spanning 17 factor-analytically derived subscales to cover mental health symptoms and psychological well-being (Watson et al., 2007; Watson et al., 2012). We administered 11 of those subscales: Appetite Loss (3 items), Dysphoria (10 items), Well-Being (8 items), Lassitude (6 items), Panic (8 items), Insomnia (6 items), Ill Temper (5 items), Suicidality (6 items), Social Anxiety (6 items), Mania (5 items), General Depression (20 items, overlapping with several of the other subscales). Participants were asked to respond to these items based on the past two weeks using a 5-point scale (1 = not at all; 5 = extremely). Prior research indicates that the IDAS subscales are internally consistent and possess good convergent, discriminant, criterion, and incremental validity (Stasik-O'Brien et al., 2019; Watson et al., 2007; Watson et al., 2008; Watson et al., 2012; Weitzner et al., 2019). Here, all of the subscales demonstrated good internal consistency (Cronbach's αs between .80 and .87) except for Insomnia and Social Anxiety, which demonstrated acceptable internal consistency (Cronbach's

[4] Of those designated to receive the intimate partner violence questionnaire, only one participant reported feeling unsafe to proceed.

αs = .77), and Mania, which demonstrated unacceptable internal consistency (Cronbach's α = .59) and so was excluded from all subsequent analyses.

Participants also completed the 4-item Perceived Stress Scale (PSS-4; Cohen, Kamarck, & Mermelstein, 1983). The PSS is the most commonly used measure of subjective stress. Exemplar items include: "In the last month, how often have you felt that you were unable to control the important things in your life?" and "In the last month, how often have you found that you could not cope with all the things you had to do?," Items are rated on a frequency scale (1 = never; 5 = very often). The PSS-4 showed acceptable reliability in this sample (Cronbach's α = .79).

Intrapersonal emotion regulation. To assess *intra*personal emotion regulation tendencies, participants completed two questionnaires at baseline: the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004) and the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003).

***Emotion dysregulation*.** The DERS is the most widely used measure of trait-like emotion dysregulation. The DERS consists of 36 items spanning six subscales, which reflect difficulties accepting negative emotions (nonacceptance), engaging in goal-directed behavior when experiencing negative emotions (goals), difficulties controlling impulses when experiencing negative emotions (impulse), lack of emotional awareness (awareness), lack of emotional clarity (clarity), and having limited access to affect-improving emotion regulation strategies (strategies). Participants were asked to indicate how often they experience these difficulties on a five-point scale (1 = almost never [0-10%]; 5 = almost always [91-100%]). In line with past research using the DERS, a total score was computed, which showed excellent internal consistency (Cronbach's α = .94).

***Cognitive reappraisal and expressive suppression*.** The ERQ is the most widely used measure of use of two specific emotion regulations strategies: cognitive reappraisal and expressive suppression. Cognitive reappraisal has been characterized as an antecedent-focused strategy that involves a person changing how they are thinking about or construing an emotion-eliciting situation or stimulus. Expressive suppression, in contrast, has been described as a response-focused strategy that involves a person inhibiting outward displays of emotion. Cognitive reappraisal and expressive suppression have been studied extensively, and both have been shown to modify emotional outcomes (Webb, Miles, & Sheeran, 2012); broadly speaking, though, cognitive reappraisal is generally thought to be more conducive to individual well-being and health than expressive suppression in the long-run (see Aldao, Nolen-Hoeksema, & Schweizer, 2010). The ERQ consists of 6 reappraisal items and 4 suppression items rated on a Likert scale (1 = strongly disagree; 7 = strongly agree). Internal consistency was good for Cognitive Reappraisal and Expressive Suppression (Cronbach's αs = .87 and .80, respectively).

Demographics, confounds, and moderators. Participants completed a demographic and relationship history questionnaire at baseline that covered gender, ethnicity, age, years of education, current employment status, the MacArthur Ladder of subjective socioeconomic status (Operario, Adler, & Williams, 2004), relationship duration, marital status, and number of children. To assess potentially important confounds and moderators, the baseline battery included self-report measures of several dimensions of personality (i.e., extraversion, agreeableness, neuroticism, emotional lability, suspiciousness, callousness, and submissiveness; Soto & John, 2016; Krueger et al., 2013), general self-efficacy (Schwarzer & Jerusalem, 1995), various interpersonal difficulties (Barkham, Hardy, & Startup, 1996), positively and negatively valenced emotional expressivity (Gross & John, 1997), perceived availability of emotional

support (Cella et al., 2007), ideal affect (Tsai, Knutson, & Fung, 2006), adult attachment (Fraley, Waller & Brennan, 2000), beliefs about emotions (Tamir, John, Srivastava & Gross, 2007), fear of negative evaluation (Leary, 1983), and external shame-proneness (Matos et al., 2015).

Other baseline measures. Several other measures were included in the baseline battery as part of ongoing collaborations. These included several questionnaires, specifically including measures of emotion-related and non-emotional related impulsivity (Carver et al., 2011), perceptions of responses to capitalization attempts (Gable, Reis, Impett, & Asher, 2004), and sleep quality (Buysse et al., 1989). They also included a 3-back working memory task (Kirchner, 1958) and an empathic accuracy task (Zaki, Bolger, & Ochsner, 2008).

Daily and Ecological Momentary Assessment

On the same days that they wore the EAR, participants also completed EMA and nightly diary surveys designed to tap into many of same constructs that were measured at baseline, as well as momentary affect.

Momentary affect. At each EMA timepoint (including the nightly diary), participants rated their current affective state. The specific affective states selected for inclusion were designed to cover each quadrant of the affective circumplex (i.e., high arousal positive, low arousal positive, high arousal negative, and low arousal negative; Feldman Barrett, 2004; Russell, 1980), as well as to cover several self-regard (e.g., pride, shame) and relational (e.g., jealous, lonely) affective states. More specifically, participants rated each of the following affective states on an intensity scale (1 = not at all; 7 = extremely): “sad/unhappy”; “anxious/worried”; “afraid/scared”; “angry/frustrated”; “ashamed”; “guilty”; “embarrassed”; “jealous”; “lonely”; “proud”; “happy”; “enthusiastic/excited”; contented/satisfied”; “confident”; “calm/peaceful”; “dull/sluggish.” Composites were created for overall positive affect (Ω_{within} = .82, $\Omega_{between}$ = .93) and negative affect (Ω_{within} = .78, $\Omega_{between}$ = .93), which were strongly negative correlated (repeated measures r = -.54).

Interpersonal emotion regulation. At each EMA timepoint (excluding the nightly diaries), participants indicated if they had wanted, asked for, or received IER from or provided IER to their partner since the last timepoint, using the same definition as above. Conditional logic was used to selectively ask follow-up questions as detailed below.

Wanting and receiving IER. If participants indicated that they had *wanted* IER (dummy-coded as 0 = no; 1 = yes), they were asked to indicate the extent to which they had felt ashamed as a result of wanting IER (1 = not at all ashamed; 7 = very ashamed) and whether they had asked their partner for IER (0 = no, 1 = yes). Regardless of whether they reported having wanted IER, participants were then asked to indicate whether their partner had tried to provide IER (0 = no, 1 = yes).

If participants reported that their partner had tried to provide IER, they were asked additional questions about that interaction, including their perceptions of what their partner did and the outcomes of that interaction. First, participants were asked to indicate the degree to which their partner: was physically present (“Was your partner present for you; did they communicate that they were available for you in the moment; did they offer a comforting gesture (e.g., a pat on the shoulder, a smile, a hug)?”), was hostile (“Was your partner harsh, critical, or dismissive”), was responsive (“Did your partner really listen to what you were saying; try to understand or empathize with your feelings; validate you?”), provided cognitive support (“Did your partner provide practical support, such as factual information, problem-solving, or advice; offer an alternative opinion or perspective?”), provided distraction (“Did your partner help you get out of your head; shift your attention away from your feelings or the situation; encourage you

to engage in pleasant or relaxing activities (e.g., go for a walk, listen to music)?"), or expressed confidence ("Did your partner express confidence in your abilities to handle the situation or your feelings?"). The items related to physical presence, hostility, responsiveness, and cognitive support were adapted from the IRIS, and we have previously shown that they correlate with responses on the IRIS (Swerdlow & Johnson, 2020). Each of these items was rated on a seven-point scale (1 = not at all; 7 = very much so).

Then, participants were asked about the outcomes of having received IER. More specifically, participants were asked about how the interaction had affected 1) how they felt overall, 2) how they felt about themselves, 3) how connected they felt to their partner, 4) their sense of control over their emotions, and 5) their ability to cope with the situation. These five items were rated on bipolar scales (-7 = felt much worse/less connected/less in control/less able to cope; 7 = felt much better/more connected/more in control/better able to cope). We have previously referred to these collectively as "perceived benefits" of receiving IER and shown that they are highly correlated at the level of the interaction (Swerdlow & Johnson, 2020); in the current data, these items were all highly inter-correlated (all repeated measures *r*s > .59), they were averaged to create a perceived benefits composite (Ω_{within} = .90, $\Omega_{between}$ = .95). Finally, participants were asked to indicate whether they had felt ashamed as a result of receiving IER (1 = not at all ashamed; 7 = very ashamed).

Providing IER. At each timepoint, participants were asked if their partner had wanted IER, if their partner had asked for IER, and if they had tried to provide IER to their partner (all dummy-coded as 0 = no, 1 = yes). When participants indicated that they had tried to provide IER, they were asked follow-up questions that paralleled the questions about receiving IER. That is, participants were asked to indicate whether they: were physically present, were hostile, were responsive, provided cognitive support, provided distraction, or expressed confidence (1 = not at all; 7 = very much so). Participants were then asked to rate their overall effectiveness in providing IER (-7 = not at all effective; 7 = very effective). Finally, participants were asked about several outcomes of having provided IER, which were written to mirror the perceived benefits of receipt. More specifically, participants indicated how providing IER affected: 1) their own emotions, 2) how they felt about themselves, 3) how connected they felt to their partner, and 4) their energy level. Each of these items was rated on a bipolar scale (-7 = felt much worse/less connected, more tired, 7 = felt much better/more connected/more energetic). The items related to how providers' own emotions, feelings about themselves, and energy were moderately correlated with each other (repeated measures *r*s between .46 and .68) and with providers' ratings of their overall effectiveness in providing IER (repeated measures *r*s between .41 and .55); however, none of these, including overall effectiveness, were correlated with the item related to providers' feelings of connection with their partners (repeated measures *r*s between -.03 and .02). As above, these items were averaged to form a perceived benefits composite (Ω_{within} = .81, $\Omega_{between}$ = .93).

Daily assessment of relationship quality. Participants answered questions about daily relationship quality and partner responsiveness as part of the nightly diary. More specifically, participants were asked to respond to the following prompts: "How well did you and your partner get along today?"; "In general, how satisfied were you with your relationship today?"; "How close did you feel to your partner today?"; "How would you rate your communication with your partner today?"; "My partner met my needs today."; "I felt understood by my partner today"; "I felt validated by my partner today"; "My partner supported me today." Each of these items was rated on a seven-point scale. These eight items were highly inter-correlated (all

repeated measures *r*s > .69); as such, they were averaged to create a composite measure of daily perceived relationship quality (Ω_{within} = .96, $\Omega_{between}$ = .99).

Daily assessment of psychological well-being and perceived stress. Participants answered questions about their subjective well-being and ability to cope with stress during the past day as part of the nightly diary. More specifically, participants responded to three items designed to tap into their daily well-being ("All things considered, how would you rate your well-being during the past day?"; "What kind of a day did you have?"; and "How optimistic are you feeling about the future?") and four items adapted from the PSS-4 to tap into perceived stress ("Did you feel that you were able to control the important things in your life today?"; "Did you feel confident about your ability to handle your personal problems today?"; "Did you find that you could cope with all the things that you had to do today?"; "Did you feel as though difficulties were piling up so high that you could not overcome them?"; adapted from Cohen, Kamarck, & Mermelstein, 1983). All seven of these items were highly inter-correlated (all repeated measures *r*s between .49 and .83), so they were averaged to form a daily stress composite, with higher scores reflecting greater stress and lower well-being (Ω_{within} = .89, $\Omega_{between}$ = .91).

Other EMA probes. As part of an ongoing collaboration, each of the EMA surveys (excluding the nightly diaries) also included a series of probes designed to assess participants' perceptions of any attempted capitalization interactions they had had with their partner (adapted from Gable et al., 2004).

End of Study

After the eight days of ecological assessment, participants completed an end-of-study questionnaire battery. This battery included several of the same questionnaires that participants completed at baseline, those being: measures of relationship satisfaction (Funk & Rogge, 2007); dyadic coping (Bodenmann, 2008), mental health and well-being (Watson et al., 2012), and perceived availability of emotional support (Cella et al., 2007).

Experiences of the EAR. Participants also completed a 10-item EAR Evaluation Questionnaire designed to assess their experience of wearing the EAR as part of the end-of-study battery (Robbins, Wright, Karan, & Baranski, 2021). Five of these items measure how obtrusive the EAR was for the participant (e.g., "To what extent did you feel uncomfortable wearing the EAR?"; 1 = not at all, 5 = a great deal). Three items measure the perceived obtrusiveness of the EAR for other people (e.g., "To what extent did the EAR influence the behavior of people around you?"). Participants were also asked to report the percentage of their waking hours that they carried the EAR on their immediate person (a slider from 0% to 100%) and the extent to which the past month had been "typical" for them (1 = not at all typical, 5 = very typical). As one would expect, these ratings were significantly inter-correlated in many cases; for example, participants who reported that wearing the EAR substantially changed their behavior were also more liable to report having felt more uncomfortable wearing the EAR (*r* = .50) and more impeded by the EAR (*r* = .53). Therefore, following past EAR research (Robbins, Wright, Karan, & Baranski, 2021), we created two subscales scores by averaging the items related to self- and other-obtrusiveness, respectively. These scores showed good internal consistency (Cronbach's αs = .84 and .82, respectively).

Audio Data Processing & Coding

I developed a coding manual expressly for coding romantic couples' IER interactions from their ambulatory acoustic data (see Appendix A). Undergraduate and post-baccalaureate content coders were trained on this manual to identify and rate the content of audibly observed

instances of IER. These trainings consisted of both live didactic instruction provided by the author and coding practice recordings, followed by review and feedback. In addition to the coding manual, coders also read and discussed an introduction to cultural informant coding adapted from the manual for the Specific Affect Coding System as part of their training (Coan & Gottman, 2007). Coders who were unable to attain adequate inadequate reliability on the practice recordings after review and feedback were delegated alternative study responsibilities. When the first coders were trained, their feedback was solicited and several improvements were made to the coding manual and training procedures to enhance clarity and precision. Throughout the coding process, all coders participated in regular team reliability meetings where they submitted ratings for randomly selected tapes to quantify inter-rater reliability and reduce rater drift.

Participants' audio files were uploaded from their recorders to a secure cloud-based server after participants returned the recorders. As detailed in the coding manual, data coding consisted of a two-phase, multiple-pass procedure. During the first pass, a coder sampled each of the audio recordings by reviewing at least the first minute of every ten-minute recording. The primary goals of the first-pass were to ascertain whether there was an audible interaction between the couple and whether that interaction was potentially IER-relevant (i.e., containing an audible verbal or nonverbal expression of emotion with both partners present) to isolate the phenomenon of interest. Coders had access to speech samples collected from participants at intake to assist with identifying participants' voices on the recordings. When *any* human speech was detected in the first minute of the recording, first-pass coders were instructed to listen to as much of that recording (and potentially the prior or subsequent recordings) as necessary to determine the presence vs. absence of an IER-relevant interaction and, if an IER-relevant interaction was present, to determine the start time and end time of that interaction so that it could be submitted for further coding.

The primary rationale for employing a periodic sampling strategy in the first-pass coding was pragmatic: considering the volume of audio recordings that we planned to collect (up to 80 hours per participant), it was deemed infeasible to review every recording in its entirety. In this regard, the approach that we used resembles that of many other EAR studies, which have traditionally relied on thin-slice sampling (e.g., recording 30-50 seconds of every 9-12.5 minutes in the case of Robbins & Karan, 2020), with the assumption that the content of these slices will substantively generalize to unobserved behavior. The key difference was that we applied our sampling strategy after-the-fact to continuously recorded audio data (opposed to only recording periodically), which allowed us to retain access to the full duration and acoustic content of the IER-relevant interactions that we observed.

During the second pass, pairs of coders independently rated the content of each interaction, including determining whether regulation transpired, and recorded their ratings in a Qualtrics survey. In some cases, a third coder also provided ratings to resolve rating discrepancies. A subset of interactions were rated as "mutually IER-relevant," meaning that there were opportunities both members of the couple to provide regulation. In these cases, at least four second-pass ratings were submitted (i.e., two from the perspective of Partner A as the would-be provider and two from the perspective of Partner B as the would-be provider). A major focus of second-pass codes was rating the emotional content of the interaction (e.g., specific emotion, emotional intensity) and the providers' enacted behavior (e.g., providing cognitive support, providing distractions). Some of these ratings were categorical (e.g., a dichotomous rating of presence versus absence of regulation), whereas others were ordinal (e.g., emotional intensity).

See Appendix A for a complete description of the coding manual, including all relevant procedures and variables.

Data Analysis Plan

I first examined univariate and bivariate descriptive statistics and visualizations for the baseline variables to characterize the sample, evaluate modeling assumptions, and consider associations between key baseline variables.

To address the first aim (Feasibility), I examined self-reported adherence and attitudes toward the EAR, as measured by an end-of-study questionnaire, and then computed zero-order correlations to identify baseline variables that predicted participants' evaluations of the obtrusiveness of the EAR. I also calculated the number of participants who requested to review their recordings and/or who asked us to delete one or more of their recordings, as well as the number of recordings that each couple produced, as complementary behavioral measures of attitudes and adherence. As before, I attempted to identify baseline variables that predicted of the number of recordings that each couple produced, using generalized linear models to model this count variable.

To address the second aim (Reliability), I first examined univariate statistics for each of the EAR codes, including identifying codes for which variability was too limited to yield interpretable metrics of inter-rater reliability (e.g., because a behavior was too infrequently observed). Then, I computed inter-rater reliability. More specifically, I computed intra-class correlations as a measure of consistency for the ordinal variables[5] and percent agreement and Krippendorff's alpha as measures of agreement for the nominal variables (Hayes & Krippendorff, 2007). Then, I calculated the within-person correlations between the second-pass codes (Bakdash and Marusich, 2017). I had originally planned to conduct a two-level exploratory factor analysis of the provider behavior codes with interactions nested within participants to evaluate the structure of the ratings (Longford & Muthén, 1992); after examining the bivariate correlations, though, I opted not to proceed with this analysis given the very low correlations.

To address the third aim (Validity), I examined correspondences between participant self-report and EAR codes. First, I computed the correlation between the proportion of a couple's EMA surveys in which they reported IER and the proportion of that couple's audio recordings that were coded as containing IER-relevant interactions. As a finer-grained test, I examined time-matched records to determine how often we did versus did not identify an IER interaction in the audio recordings when participants did versus did not report having engaged in IER in that same epoch. The purpose of these analyses was to ascertain how often we might be detecting IER interactions that were "invisible" to the participants, as well as how often we might be missing IER-relevant interactions. As an exploratory test of what makes IER interactions more or less visible, I contrasted the IER interactions for which there was apparent consistency between the EAR and EMA records (i.e., we identified an IER interaction with the EAR and one or both participants reported having engaged in IER during that same epoch) with those for which there was apparent inconsistency (i.e., participants reported having engaged in IER, but we did not observe an IER interaction with the EAR during that epoch). To do this, I computed a series of parallel random-intercept mixed effects logistic regression models (interactions nested within dyads).

---

[5]Alternative measures of consistency, including the concordance correlation coefficient (Lin, 2000), Kendall's *W*, Pearson's *r*, and Spearman's *rho* yielded substantively identical results. As such, we only report the intra-class correlations here.

Then, I examined self-report correlates and predictors of our provider behavior codes. First, I examined correspondence between a subset of the provider codes and participants' self-reports of having received or provided of those same behaviors in their EMA surveys (e.g., coder and participant ratings of the degree to which the provider exhibited hostility). I did this in two ways: first using time-linked records (within-person correlation) and second using aggregates (zero-order correlation). The rationale for these parallel analyses was to consider the possibility that participants' ratings might sometimes be referencing different interactions than coders' ratings, in which case the zero-order correlations between the aggregates might actually be expected to be stronger than the within-person correlations. Next, I examined whether the provider codes could be predicted from key baseline measures. Finally, I tested the *a priori* hypothesis that the frequency of IER interactions would be associated with the interaction of perceived responsiveness and stress. Again, I did this in two ways: first daily ratings of relationship quality and stress (mixed effects regression) and second using baseline measures of relationship quality and stress (OLS regression). Here, the purpose of the analyses was to contrast contemporaneous analyses (i.e., the association between IER frequency and same-day perceived stress) with time-lagged analyses (i.e., the association between IER frequency and baseline perceived stress).

All analyses were conducted in R version 4.0.2 (R Core Team, 2020). All reported regression coefficients are standardized. Repeated measures variables were centered and standardized within-person prior to analyses (Wang et al., 2019), and we followed best practice recommendations for conducting mixed effects modeling with a modest level-2 sample, including using restricted maximum likelihood and the Kenward-Roger approximation (McNeish, 2017). Where data generation processes yielded variables that were significantly non-normal (e.g., count variables), we computed nonparametric rank-order associations and generalized linear models. Multilevel reliability was computed using the multilevelTools package (Wiley, 2020). Repeated measures correlations were computed using the rmcorr package (Bakdash & Marusich, 2017). Ordinal regressions and negative binomial regressions were computed using the ordinal and MASS packages, respectively (Christensen, 2019; Venables & Ripley, 2002). Mixed effects modeling was implemented using the lme4 (Bates, Machler, Bolker, & Walker, 2015), lmerTest (Kuznetsova, Brokhoff, & Christensen, 2017), and pbkrtest (Halekoh & Højsgaard, 2014) packages. Inter-rater reliability metrics were computed with the irr package (Gamer, Lemon, Fellows, & Puspendra, 2012). Intraclass correlations were calculated using the sjstats package (Lüdecke, 2021).

Results

Preliminary Analyses

Descriptive statistics for all of the key baseline variables, including skewness, kurtosis, and within-dyad correlations, are reported in Table 1.

Sample Representativeness

Sample representativeness was considered along several axes. In terms of ethnicities, the sample was roughly in line with 2010 census data for the San Francisco Bay Area. The average participant placed themselves just above the midpoint on the MacArthur ladder of subjective socioeconomic status "in their community" (M = 5.87, SD = 2.01).

Overall, participants in our sample reported being highly committed to their relationships (M = 9.47 out of 10, SD = 1.22), consistent with our intention to recruit couples in committed romantic relationships. Only 20 participants (16.94%) provided CSI-16 ratings indicative of notable relationship dissatisfaction (i.e., CSI-16 < 51.5; Funk & Rogge, 2007), resulting in a

moderately left-skewed distribution of relationship satisfaction scores. Notably, though, of the subset who answered the intimate partner violence questionnaire, nearly 45% reported 1+ instances of intimate partner violence in the past year. The distribution of intimate partner violence scores was highly right-skewed and positively kurtotic.

In terms of mental health, symptom scores and well-being scores were broadly in line with national norms for the IDAS-II (Nelson, O'Hara, & Watson, 2018). Between 7% (Ill Temper) and 32% (Lassitude) of participants reported mild or greater symptoms on each of the 11 administered subscales, but very few participants reported severe symptoms (< 3% of participants for all subscales) (based on clinical severity cut-offs calculated by Stasik-O'Brien et al., 2019). Applying stringent diagnostic prediction cut-offs, 19% of participants in the sample would likely have met criteria for an internalizing diagnosis based on their baseline symptom report (Stasik-O'Brien et al., 2019).

Correlations between Key Baseline Variables

As shown in Table 1, strong correlations between partners' ratings (i.e., within-couple consistency) were observed for scores on the CSI-16, PPRS, and DCI, such that if one partner provided relatively high ratings of relationship satisfaction on the CSI-16, their partner also tended to provide higher ratings of relationship satisfaction. More modest within-dyad correlations were observed for the other key baseline variables, including the RLS and the IDAS subscales.

Not surprisingly, substantial correlations between the baseline measures of relationship quality were observed (all *r*s between .43 and .73), such that participants who tended to provide higher ratings of relationship satisfaction on the CSI-16 also tended to provide higher ratings of perceived partner responsiveness on the RLS, PPRS, and DCI Evaluations subscale. The IDAS-II subscales were likewise substantially inter-correlated (median $r = |.44|$; see also Watson et al., 2012 and Wester at al., 2021 for prior factor analytic evidence that these IDAS-II subscales load onto a higher-order factor). Therefore, to reduce dimensionality in subsequent analyses, I created composites of Relationship Quality and Internalizing Symptoms variables. To do this, I Z-scored participants' scores on the CSI-16, PPRS, RLS, and DCI (Relationship Quality) and the subscales of the IDAS-II (excluding the General Depression subscale, which contains items that overlap with the other subscales; Internalizing Symptoms) and then averaged these standardized scores (Cronbach's αs = .85 and .86 for Relationship Quality and Internalizing Symptoms, respectively).

Bivariate correlations between these composites and other key baseline variables are reported in Table 2. As shown, participants who reported higher composite relationship quality tended to report lower intimate partner violence, internalizing symptoms, perceived stress, and use of expressive suppression. Those who reported higher composite internalizing symptoms tended to report higher perceived stress, emotion dysregulation, and expressive suppression. Perceived stress was likewise positively correlated with emotion dysregulation and expressive suppression. Emotion dysregulation was positively correlated with expressive suppression and negatively correlated with cognitive reappraisal use.

Aim 1: Feasibility

To consider the feasibility, I examined both participant self-reports regarding their experience of the EAR and the characteristics of the acquired EAR recordings.

Participant Experiences of Wearing the EAR

Of the couples who dropped out of the study after consenting, only one couple cited discomfort with the EAR as their reason for dropping out. In each nightly diary, participants had

the opportunity to designate any segments of time that they would wish us to delete before data processing. Out of all completed nightly diaries (k = 635), fewer than 10% contained requests to delete segments of any length, and the median request was to delete a single segment of less than one hour in duration (out of 10 hours of total recording per day). In a similar vein, only one couple requested to review any of their recordings before data processing despite this option being communicated to all couples.

The median participant reported that they had carried the EAR on their person 80% of the time on their designated days of the week (M = 78%; SD = 18.34%). On average, participants reporting being somewhat aware of wearing the EAR (M = 3.17; SD = .91), a little uncomfortable wearing the EAR (M = 2.20; SD = 1.02), having been a little impeded in their daily activities by the EAR (M = 1.70; SD = .83), having changed their way of talking a little because of the EAR (M = 1.89, SD = .95), and having changed their behavior a little because of the EAR (M = 2.00, SD = .83). Whereas participants reported that they had talked to people around them about the EAR somewhat (M = 2.83, SD = 1.21) and that people around them were somewhat aware of the EAR (M = 2.91, SD = 1.30), they also indicate that the EAR had not had a strong influence on the behavior of people around them (M = 1.74, SD = 1.04).

Participants' ratings of the obtrusiveness of the EAR for themselves were moderately correlated with their ratings of the obtrusiveness of the EAR for others (r = .28, p = .003). Ratings of both these types of obtrusiveness were associated with how much of the time participants reported wearing the EAR, such that participants who regarded the EAR as more intrusive tended to have worn the device less of the time (r = -.25, p = .007 and r = -.24, p = .01). Across couples, partners' ratings of their experiences of wearing the EAR were generally moderately correlated, such that substantial within-couple consistency was observed for ratings of obtrusiveness for themselves (r = .36) for others (r = .38). In other words, if a participant reported that wearing the EAR had been relatively obtrusive, then their partner was also likely to report that the EAR was more obtrusive.

Next, we sought to identify baseline variables that were associated with ratings of obtrusiveness. As shown in Table 3, several demographic variables were associated with self-reported obtrusiveness; more specifically, female participants (vs. male participants), older participants, participants with more years of education, participants with longer relationship durations, and participants with children were more liable to rate the EAR as relatively obtrusive for themselves—although none of these demographic variables were related to ratings of obtrusiveness for others. More frequent use of reappraisal predicted lower self-obtrusiveness. More frequent intimate partner violence predicted lower other-obtrusiveness. No significant associations were observed between ratings of obtrusiveness and the relationship quality composite, the internalizing symptom composite, perceived stress, emotion dysregulation, or use of expressive suppression.

Multiple personality variables predicted participants' evaluations of wearing the EAR, as shown in Table 3. More specifically, higher scores on measures of fear of negative evaluation, external shame-proneness, and interpersonal problems and lower levels of positive expressivity, extraversion, and self-efficacy were associated with higher reports of self-obtrusiveness. In a similar vein, higher levels of fear of negative evaluation, attachment insecurity, and interpersonal problems and lower levels of openness with others and self-efficacy were associated with higher reports of other-obtrusiveness. In summary, participants' ratings of their experience of wearing the EAR were systematically associated with a number of demographic and personality variables, but few of the key variables of the present investigation.

EAR Data Acquisition

Participants produced 13341 recordings that were submitted for data processing (approximately 2200 hours of audio), or an average of 128 recordings per participant. As both members of each couple were wearing recorders, though, many these recordings were essentially duplicative for our purposes in that they captured the same time intervals. As such, the first coding pass ultimately consisted of reviewing 8197 recordings.[6] Of these, the couple were observed to be physically present or interacting with each other in roughly a third (2555 recordings), and roughly a third of those (836 recordings) were designated as potentially IER-relevant (i.e., containing an audible verbal or nonverbal expression of emotion with both partners present) and submitted for more detailed second-pass coding.

The interactions that were classified as potentially IER-relevant and submitted for second-pass coding were split roughly evenly by the valence of the primary emotion to which the would-be provider was responding (52.1% negative valence, 47.9% positive valence). Among the negatively valenced emotions, anger/frustration was by far the most common (28.3% of all IER-relevant interactions), followed by anxiety/fear and sadness. About a third of the IER-relevant interactions were classified as mutually IER-relevant (31.2%), meaning that there were potential opportunities on both sides of the couple for IER. We captured a variety of types of emotion-eliciting situations, including those that were internal to the relationship and where the cause of the receivers' emotions was the provider (39.2%); those that were internal to the relationship, but the cause of the receivers' emotions was not the provider (20.7%); and those that were external to the relationship (40%). Of note, most of the IER-relevant interactions were quite brief, averaging only about 2 minutes in duration from the start of the interaction to the end of the interaction, but there was considerable variability in duration across interactions (SD = 190.96 seconds).

There was substantial variability in the number of recordings that participants produced (SD = 141.53). Moreover, across couples, there was a strong correlation between the number of recordings produced by one participant and the number produced by their partner (r = .79), so that if a participant produced a relatively large number of recordings, then their partner also tended to have produced more recordings. This variability across participants and couples can be attributed most directly to a combination of participant nonadherence (e.g., not recharging or turning on the EAR) and technological failures (e.g., file corruption).[7]

---

[6] "Duplicative" recordings were used by coders on a supplemental, as-needed basis when the content of a recording could not be thoroughly assessed from one partner's recording of that epoch alone (e.g., when the interaction consisted of the partners speaking with each other on the phone).

[7] Early in the data collection phase, several couples reported to the researchers that their EAR devices had apparently crashed due to the phone storage being full. Subsequent testing revealed that some of the EAR devices were not properly transferring the audio recordings to the microSD card, which resulted in the devices' internal storage capacities eventually being exceeded. We eventually discovered that this was due to the devices' firmware configurations; in some cases, though, we were unable to update the firmware due to the age of those devices. Considering that the currently supported EAR form factor consists of an Android app that is under active development—and that many dozens of Android devices are available—this experience underscores the importance of thorough testing of all aspects of device compatibility.

I attempted to identify variables that were associated with the number of EAR recordings obtained per participant. The count of recordings was overdispersed relative to a Poisson distribution. Accordingly, I computed a series of parallel negative binomial models in which number of recordings was entered as the dependent variable to examine this question. Somewhat surprisingly, no significant associations were observed between the obtrusiveness ratings and the number of recordings produced (*p*s > .26). In fact, none of the baseline demographic variables, key variables, or confounds that we considered were systematically associated with the number of recordings produced. On the other hand, the date when a couple was consented into the study accounted for nearly 22% of the variance in the total number of recordings produced by that couple (IRR = 1.82, $p < .001$), potentially consistent with improvement in researcher data acquisition procedures over time.

Aim 2: Reliability

Next, I considered the reliability and structure of the coding system. Several common challenges were identified by both first- and second-pass coders. These were: identifying the participants from their voices, background noise, the presence of speakers other than the participants, poor audio quality (e.g., voices muffled, low volume). First-pass coders also identified determining the start-time and end-time of potentially IER-relevant interactions as a common challenge, whereas second-pass coders identified very brief interactions (e.g., interactions lasting only several seconds) and lack of context. The second-pass codes that were most commonly endorsed as challenging to rate were: Locus of the Interaction, Primary Emotion, Provider and Receiver Satisfaction, Attunement, and Provider Expressions of Acceptance.

A key question was whether we would be able to reliably distinguish potentially IER-relevant recordings from non-relevant recordings, as this first-pass determination was the funnel to the more detailed second-pass coding. Reliability for this code was ascertained from team reliability meetings in which all first-pass coders submitted ratings for batches of randomly selected recordings. Of the 22 research assistants who underwent first-pass coding training, all of them attained adequate reliability on training tapes. Inter-rater agreement was very good for coding IER-relevance (% Agreement = .87, Krippendorff's alpha = .84).

Each potentially IER-relevant interaction was coded by two or in some cases three (i.e., to resolve categorical disputes) independent coders, and these sets of ratings were used to calculate metrics of inter-rater reliability. Of the 15 research assistants who underwent second-pass coding training, two (13.33%) failed to attain adequate reliability on training tapes after feedback. Six of the second pass codes demonstrated inadequate variability to be included in further analysis due to low base rates (i.e., they produced uninterpretable inter-rater reliability metrics). These were: the manner in which IER was solicited by the would-be receiver, whether IER was offered by the would-be provider, and whether IER was accepted by the would-be receiver, as well as three provider behaviors: Expressions of Confidence, Expressions of Overwhelm, and Expressions of Acceptance. As shown in Table 4, substantial inter-rater consistency and agreement was observed for all of the remaining second-pass codes, although the strength of the reliability ranged from excellent to tolerable depending on the specific code in question (ICCs ranging from .51 to .92, Krippendorff's alphas ranging from .52 to .70). Disputes between second-pass coders were resolved before subsequent analyses by taking the average between the coders in the case of ordinal variables and by a tie-breaking third coder in the case of nominal variables.

As shown in Table 5, ratings of the provider behavior codes showed modest-to-moderate clustering (interactions nested within participants) such that the preponderance of the variance was within-person (opposed to between-persons) for all of the provider behaviors with the exception of Expressive Suppression. In addition, all of these variables were moderately-to-highly right-skewed and kurtotic, which is unsurprising given that coders were instructed to base their ratings primarily on the number of observed utterances that matched each provider behavior, resulting in a high incidence of zeroes.

As shown in Table 6, statistically significant within-person associations were observed between many of the provider behaviors. In general, though, these associations were of quite modest magnitude (median $r = |.08|$). Contrary to observations made in developing and validating the self-report Interpersonal Regulation Interaction Scale, Cognitive Support and Practical Support were not significantly correlated and Empathic Responding and Harsh Responding were only modestly negatively correlated, suggesting that these behaviors may be more distinctive than participant retrospective self-report might indicate. The largest correlations that were observed were for four codes that capture what might be regarded as more negatively valenced provider behaviors, those being: Harsh Responding, Defensiveness, Expressive Suppression, and Co-Brooding, which were moderately inter-correlated (*r*s ranging from .15 to .28), suggesting that these behaviors frequently co-occur within interactions. No other correlations between provider behaviors exceeded $r = |.15|$. Comparable correlation coefficients were obtained when only the subset of interactions related to negatively valenced emotions was examined. I had originally planned to submit these ratings to a two-level exploratory factor analysis to further explore their structure; in light, though, of these modest associations and associated communalities, I opted not to proceed with this step.

The provider behaviors were also associated with other features of the interaction that we coded. Not surprisingly, most of the provider behaviors that we coded occurred at significantly higher rates when the primary emotion was negatively valenced (versus positively valenced), as shown in Table 7. The exceptions were Expressions of Warmth, which occurred significantly more when the primary emotion was positively valenced, and Empathic Responding, for which the association with emotional valence was not statistically significant. Given the low base rates for some of these key codes in positively valenced interactions and to eliminate valence as an interpretive confound, I narrowed in on the subset of negatively valenced interactions. In this context, some provider behaviors occurred at significantly higher rates when the cause of the receiver's emotion was internal to the relationship (vs. external to the relationship), those being: Harsh Responding, Reassurance, Expressions of Warmth, and Defensiveness. Conversely, Empathic Responding occurred at a significantly *lower* rate when the locus of the interaction was internal to the relationship. Within the set of internal locus interactions, interactions where the provider themself was the cause of the receiver's emotion were associated with significantly higher rates of Harsh Responding, Defensiveness, and Co-Brooding and lower rates of Empathic Responding and Distraction. In a similar vein, interactions that were coded as mutually IER-relevant were rated as significantly higher in Harsh Responding, Defensiveness, Expressive Suppression, and Co-Brooding. Higher receiver emotional intensity at the start of the interaction was associated with greater Harsh Responding, Reassurance, Defensiveness, Expressive Suppression, and Co-Brooding, as shown in Table 8. Greater use of Empathic Responding and Distraction were unique in that they were the only behaviors that were significantly associated with change in the receiver's emotional intensity from the start of the interaction to the end of the interaction, such that greater use of these strategies was associated with reductions in emotional

intensity, on average. Receivers and providers both tended to be coded as more satisfied when the provider was more empathic and warmer, less harsh and defensive, offered more opportunities for distractions, and engaged in less expressive suppression and co-brooding.

Aim 3: Validity

To address my third aim, I conducted a series of analyses to examine the validity of the coding system with a focus on correspondences between the EAR ratings and participant self-report. There was a positive correlation between the proportion of a couples' audio recordings that were coded as IER-relevant and the proportion of a couples' EMA records in which one or both participants reported having wanted, asked for, received, or provided IER (r = .28). Consistent with this modest correlation, we often detected IER interactions in the EAR recordings during epochs in which both members of the couple reported *not* having received or provided IER. More specifically, out of 507 IER interactions that were observed via the EAR and for which we also had time-matched EMA data available from both partners, both partners reported *not* having received or provided IER in roughly two-thirds of these cases (67.5%).

Among the IER interactions observed on the EAR, what distinguishes whether one or both participants reporting having engaged in IER during the same epoch? I conducted a series of parallel mixed effects logistic regression models (interactions nested within dyads) to examine this question. Higher coder ratings of co-brooding (OR = 1.24, p = .04) and defensiveness (OR = 1.28, p = .02) were associated with a higher likelihood of participants endorsing having provided or received IER in the same epoch as the interaction, as did higher emotional intensity at the start of the interaction (OR = 1.30, p = .01) and negative valence (vs. positive valence, OR = 1.63, p = .04). None of the other coder ratings significantly differentiated the likelihood of an accompanying report of IER from either participant, specifically including the other provider behaviors (ps > .06), the interaction locus (ps > .59), the emotional intensity at the end (p = .24), the receiver's satisfaction (p = .19), the provider's satisfaction (p = .22), or the duration of the interaction (p = .58). In other words, there was more correspondence between self-report and our EAR coding regarding the presence versus absence of IER interactions when the interactions in question involved higher intensity, negatively valenced emotions, more co-brooding, and more defensiveness.

For the epochs where we observed one or more IER interactions *and* had EMA ratings from the participants, I examined whether there was correspondence between coder ratings and participant ratings of those same epochs for a set of variables where there was a participant probe that conceptually matched one of our EAR codes, those being: Empathic Responding, Harsh Responding, Cognitive Support, Providing Distractions, Provider Satisfaction, and Receiver Satisfaction. As shown in Table 9, statistically significant positive associations were observed for ratings of Empathic Responding, Harsh Responding, Provider Satisfaction, and Receiver Satisfaction, although the magnitude of these associations was modest. Moreover, the magnitude of the correlations was substantively unchanged when we examined the subset of negatively valenced interactions in *post hoc* analyses.

As a complementary strategy for investigating correspondence between the EAR data and the EMA data, I also examined associations between aggregates of these ratings across all epochs. For example, I examined whether participants that we coded as more responsive were also rated as more responsive by their partners when averaging across all IER interactions. As shown in Table 10, significant positive associations were observed—and, qualitatively speaking, these associations tended to be larger than those observed in the time-matched analyses.

I also examined whether the aggregated EAR ratings could be predicted from baseline scores on key measures. Somewhat surprisingly, few robust links were observed between the EAR codes and baseline measures of internalizing symptoms, perceived stress, intrapersonal emotion dysregulation, and intrapersonal reappraisal and suppression, as shown in Table 11. There were several exceptions, though; namely, higher scores on measures of internalizing symptoms, perceived stress, and emotion dysregulation predicted less provision of reassurance; reappraisal use predicted less provision of practical support and more co-brooding; and suppression use predicted more provision of expressive suppression and defensiveness. The more relationship-specific baseline measures of relationship quality and intimate partner violence showed more robust correspondences with EAR codes. That is, higher baseline relationship quality (participant self-report) predicted more empathic responding, less harsh responding and defensiveness, and higher satisfaction with IER interactions (coder ratings). Conversely, more frequent experiences of intimate partner violence strongly predicted less empathic responding and warmth; more harsh responding, defensiveness, and co-brooding; and lower satisfaction with IER interactions.

Next, I examined the relationships between stress, relationship quality, and the number of IER interactions captured on the EAR. First, I examined the within-person relationships between these variables using our daily measures of stress and relationship from the nightly diaries. Significant main effects of either same-day stress (IRR = .97, p = .41) or same-day partner responsiveness (IRR = 1.03, p = .49) were not observed; however, a significant interaction between same-day stress and same-day partner responsiveness was observed (IRR = .91, p = .005). As shown in Figure 1a, a cross-over interaction was observed. On days when perceived stress was higher-than-average, higher-than-average relationship satisfaction was associated with a lower count of IER interactions. Conversely, on days when perceived stress was lower-than-average, higher-than-average relationship satisfaction was associated with a higher count of IER interactions. Further *post hoc* investigation of this interaction revealed that the interaction in question was obtained for the count of negatively valenced IER-relevant interactions (p < .001), but not for the count of positively valenced IER-relevant interactions (p = .71; see Figures 1b and 1c). As a complement to these analyses, we examined whether the proportion of a couple's recordings that contained IER interactions could be predicted from the interaction between baseline perceived stress and baseline relationship quality. As above, no significant main effect of baseline perceived stress (b = .16, p = .14) or baseline relationship quality (b = -.17, p = .13) was observed; however, a significant interaction was again observed (b = .17, p = .04). The shape of the interaction in this *between*-person analysis matched our original prediction. That is, baseline perceived stress positively predicted the rate of IER-relevant interactions for participants who reported higher levels of relationship quality at baseline and negatively predicted the rate of IER-relevant interactions for participants reported lower levels of relationship quality at baseline.

Discussion

Notwithstanding increased interest in processes of IER over the past decade, there continues to be a paucity of ecologically valid, dyadic data available to inform understanding about these dynamic, transactional processes. Therefore, the aim of the current study was to investigate the feasibility, reliability, and validity of a novel approach to studying IER interactions in daily life. More specifically, I sought to evaluate the use of ambulatory acoustic recording in tandem with ecological momentary assessment to simultaneously assess subjective experiences of and behaviors enacted during spontaneously occurring IER interactions. In the

process, I developed a coding manual for the express purpose of identifying and rating the content of acoustically observable IER interactions, which I sought to validate. To accomplish these aims, my team and I recruited and followed a community sample of committed, cohabiting adult romantic couples, which I regard as a model dyad to study interpersonal emotion regulation processes.

With regard to feasibility, most participants reported that wearing the EAR was reasonably unobtrusive, consistent with prior studies (e.g., Mehl & Holleran, 2007; Manson & Robbins, 2017). Participants rarely requested us not to listen to segments of their audio despite multiple opportunities throughout the study to make such requests, and only one couple specifically cited discomfort with the EAR as their reason for withdrawing from the study. Moderate consistency was observed in couples' ratings of their experiences of the EAR, such that if one participant rated the EAR as relatively obtrusive, their partner was liable to have also rated the EAR as relatively obtrusive. In line with this, several demographic variables were associated with participants' higher ratings of obtrusiveness, including older age, more years of education, longer relationship duration, and female gender. Likewise, a plethora of personality variables were at least moderately strongly associated with ratings of obtrusiveness, including extraversion, external shame-proneness, general interpersonal problems, and attachment insecurity, partly consistent with prior research reporting associations between the HEXACO personality traits and willingness to volunteer for research involving ambulatory acoustic recording (Manson & Robbins, 2017). On the other hand, few of the key baseline measures of relationship quality, mental health, or emotion regulation were significantly correlated with ratings of obtrusiveness.

Crucially, whereas a range of baseline variables were associated with self-reported experiences of wearing the EAR, none of those variables, including self- or other-obtrusiveness, was associated with the number of recordings that we obtained from participants. In other words, although some participants were more likely to regard the EAR as obtrusive than others as a function of demographic and personality variables, those participants were not systematically less likely to provide usable data. Somewhat surprisingly, the date that a couple entered the study was a strong predictor of the number of recordings that they produced. Although I consulted heavily with one of the developers of the EAR prior to study launch, I suspect that this association reflects systematic improvement over time in our data collection procedures, including both technical (e.g., device configuration) and non-technical (e.g., sharpening our pitch to participants of the rationale for wearing the EAR as much as possible) procedures.

Notwithstanding challenges in data acquisition, we collected nearly 2200 hours of audio across all participants. Crucially, more than one in ten of the recordings that were submitted for first-pass coding were observed to contain a potentially IER-relevant interaction between the couple despite the fact that our coders did not listen to the entirety of every one of those recordings. This observation suggests both that the daily lives of romantic couples are rife with opportunities for IER and that it is feasible to obtain a large sample of IER interactions using a few days of acoustic monitoring with the EAR. Moreover, the IER-relevant interactions that we captured spanned a wide range of experiences, including a variety of specific emotions and types of emotion-eliciting situations. Indeed, we could potentially have reduced the total number of recording days and still have collected an adequate sample of IER interactions, particularly as this would have allowed us to sample the recorded audio at a higher frequency during coding. At the same time, some of the behaviors that we coded turned out to be very rare in this sample; to

the extent that these behaviors are of interest in future studies, consideration would need to be given to larger or more targeted samples.

A critical question from the outset was whether we would be able to reliably distinguish IER-relevant interactions from non-IER-relevant interactions. To make this distinction, coders had to determine first whether both members of the couple were physically present or interacting with each other and second whether either member of the couple expressed an emotion verbally or nonverbally. Notwithstanding the multiple challenges involved in making these determinations, all of the potential coders were able to be adequately trained and inter-rater reliability was good. In fact, inter-rater reliability ranged from adequate to excellent for all of the codes in our manual, although a handful of codes showed inadequate base rates and variability to be included in these or subsequent analyses (e.g., only 2.1% of IER interactions were coded as containing any instances of the provider expressing confidence in the receiver or the relationship).

Having ascertained the reliability of our first- and second-pass codes, I next turned my attention to the structure of those codes. Following on past work developing the IRIS, I was particularly interested to examine the correlations between coder ratings of the providers' behavior. In developing the IRIS, for example, we found that receivers' retrospective ratings of the items "They helped me to solve the problem" and "They suggested alternative interpretations of the situation" were quite highly correlated—to the point that they were statistically subsumed by a the same higher-order dimension, which we have referred to as Cognitive Support (Swerdlow & Johnson, 2020). We have previously suggested that such a correlation could potentially arise from multiple processes, including people being insensitive to the conceptual distinction between problem-solving and cognitive reappraisal or these behaviors systematically co-occurring at a high rate within interactions, such that if a person receives a lot of problem-solving during a particular interaction, they are likely to also receive a lot of reappraisal.

Here, though, all of the correlations between the coder ratings of provider behaviors—including the correlation between problem-solving (Practical Support) and co-reappraisal (Cognitive Support)—were quite modest, so much so that I opted not to proceed with a planned exploratory factor analysis. This finding would tend to suggest that these behaviors are more distinctive than retrospective participant self-report might credit, although the combination of restricted variability and measurement error may have played a role in constraining the correlations. At the same time, it is intriguing that the strongest correlations were observed between the four behaviors that might broadly be regarded as relatively negatively valenced contributions to the interaction: Harsh Responding, Defensiveness, Expressive Suppression, and Co-Brooding. This would suggest that these behaviors are more liable to systematically co-occur, such that if a particular IER interaction contained one or more instances of the provider engaging in one of these behaviors, that same interaction was more likely to contain instances of the other three behaviors.

Intraclass correlations for all of the provider behaviors were likewise modest. That is, although there was evidence that each behavior was used more by some providers than others, providers were apparently quite variable in their behavior across interactions, perhaps reflecting sensitivity to the situational context (cf. Pauw, Sauter, van Kleef, & Fischer, 2019). Unsurprisingly, for example, participants were much more likely to engage in defensive behavior when the locus of the interaction was internal to the relationship and the cause of the receiver's (negative) emotion was the provider. Thus, whereas there may be considerable test-retest reliability in romantic couples' behavior during conflict conversations (Gottman & Levenson,

1999), these behaviors may not generalize to other contexts. This underscores the usefulness of an intensive repeated measures approach that allows sampling a couples' behavior across a wide swath of contexts.

Notwithstanding the considerable within-person variability, provider behavior could still be predicted from key baseline measures. The specific pattern of associations observed in these analyses was generally highly interpretable with, for example, baseline scores on a measure of use of expressive suppression positively predicting interpersonal provision of expressive suppression. Whereas both internalizing symptoms and intrapersonal emotion dysregulation are associated with more reassurance-*seeking* (e.g., Dixon-Gordon, Haliczer, Conkey, & Whalen, 2018; Fearey, Evans, & Schwartz-Mette, 2021), these same variables predicted less reassurance *provision* in our sample. Intriguingly, compared to these measures of intrapersonal tendencies and functioning, the more interpersonally oriented, relationship-specific measures of relationship quality and intimate partner violence showed a considerably more robust pattern of associations with provider behavior. In this regard, the associations with intimate partner violence were particularly striking given both the magnitude of these associations (e.g., $r = .63$ with co-brooding) and the profound consequences of intimate partner violence for personal and public health (Cambell, 2002; Smith et al., 2018). That the dynamics of IER interactions as measured by our coding system track so closely with experiences of intimate partner violence is a strong validation of both IER as a conceptual domain and the method in question, potentially opening the door to new approaches to identification, intervention, or prevention efforts.

I also examined correspondences between our coder ratings and participants' EMA data in further tests of validity. As expected, there was a significant, positive correlation between the total proportion of a couples' EAR recordings that we coded as containing an IER-relevant interaction and the total proportion of a couples' EMA responses that indicated engagement in IER, although the magnitude of this correlation was moderate. Indeed, many of the IER-relevant interactions that we identified in the first pass of coding went unreported by participants in that we observed IER-relevant interactions during epochs where both participants reported *not* having engaged in IER. This observation supports the idea that a major strength of the EAR is that it allows us to capture IER interactions that are apparently "invisible" to the participants as such. This invisibility could arise through multiple processes, including not labeling interactions as IER (based on the definition provided) or forgetting about interactions altogether. For example, although the definition of IER provided to participants was technically inclusive of situations in which the target emotion was positively valenced, agreement between participants and coders regarding the presence of IER in a given epoch was considerably higher when the target emotion was negatively valenced. In other words, the EAR may allow us to understand a richer array of IER-relevant interactions than relying solely on participants to tag their IER interactions.

There was some evidence for cross-method correspondence in the *content* of IER interactions. In time-linked analyses, statistically significant correlations were observed between coder ratings of Empathic Responding, Harsh Responding, Cognitive Support, and Satisfaction for a given interaction and participant ratings of those same variables at the corresponding EMA timepoint. These correlations, though, were notably modest in magnitude. One factor that likely contributed to the size of these correlations was that participants and coders were operating off of different—albeit conceptually overlapping—prompts and scales for rating these variables. A second complicating factor is that it is not clear whether or when participants' ratings were referencing the same interaction or a different interaction as the one that our coders or their partner were rating. That is, at each EMA timepoint, participants were asked to report on IER

interactions transpiring at any point in the past several hours, as ratings were taken four times per day. One intriguing insight from the current study was the sheer volume of brief IER interactions that were coded, with many epochs containing more than one coder-observed IER interaction. Thus, although these analyses were time-linked to a degree, participants' ratings may have frequently been referencing a different interaction, in which case we might not expect to obtain strong correlations.

In light of this, I conducted complementary analyses in which I aggregated across all EAR and EMA ratings for each participant. That is, if coders generally rated a participant as providing more cognitive support, did the participant and their partner also tend to rate them as generally providing more cognitive support (across EMA surveys)? The correlations observed in these analyses were consonant with, but qualitatively larger than those observed in the time-linked analyses—supporting the notion that one source of error in the time-linked analyses was likely that participants were sometimes referencing different interactions during those epochs than we were. A key practical goal for future studies using this method, then, will be to use a more dependable method for matching self-report and coder ratings.

Finally, I tested the pre-registered prediction that the frequency of IER interactions observed via the EAR would be highest when perceptions of relationship quality were high *and* perceptions of stress were high. I tested this in two ways. First, I examined whether the count of IER interactions varied as a function of same-day relationship quality and stress ratings, as measured by participants' nightly diary responses. Second, I examined whether the proportion of a couples' EAR recordings containing IER-relevant interactions could be predicted from scores on baseline measures of relationship quality and perceived stress. Here, a divergence was observed between the results of the former within-person analyses and the latter between-person analyses. Results of the between-person analyses closely matched hypotheses. That is, higher baseline perceived stress predicted a higher proportion of IER-relevant interactions for those who reported higher baseline relationship quality, but a lower proportion of IER-relevant interactions for those who reported lower relationship quality. This is consistent with prior research indicating that perceived responsiveness facilitates emotion expression (Ruan et al., 2020).

The within-person analyses likewise yielded a statistically significant interaction between stress and relationship quality, but with a different pattern of the association. On days when relationship quality was higher-than average, higher-than-average perceived stress was associated with a *lower* count of IER interactions, whereas on days when relationship quality was lower-than-average, higher-than-average perceived stress was associated with a higher count of IER interactions. Further investigation of this interaction indicated that this association was observed for the count of same-day negatively valenced IER interactions, but not the count of same-day positively valenced IER interaction. Speculatively, then, this within-person interaction could reflect an undermining of relationship satisfaction that arises in the context of high stress leading to multiple, perhaps ineffectual or conflictual, negatively valenced interactions with one's partner.

Strengths and Limitations

This study had several notable strengths and limitations, several of which have already been noted. My team and I collected data from a community sample that was reasonably diverse along multiple axes, including ethnicity, socioeconomic status, and age. With one exception, all study procedures and measures were completed by both members of the couple, in line with the inherently dyadic processes that we were seeking to study. In addition to an extensive baseline battery, we also collected intensive repeated measures data using two complementary,

ecologically valid approaches in EMA and ambulatory acoustic recording, which allowed us to assess both subjective experience and enacted behavior.

Although our level-1 sample sizes were adequate (e.g., k = 836 IER-relevant interactions), perhaps the most notable limitation of the present investigation is that our level-2 sample size was modest (n = 118 completers, or 59 dyads). We had initially pre-registered a plan to enroll 75 dyads into the study; this plan, though, was complicated by the onset of the COVID-19 pandemic, a contemporaneous decline in participant adherence, and other pragmatic factors. Where possible, we have used more conservative procedures for inferential statistics (e.g., using restricted maximum likelihood and the Kenward-Roger approximation when computing mixed effects models; see McNeish, 2017). Apparently null effects should be regarded with caution, though. We conducted several exploratory analyses that entailed a large number of parallel comparisons, and although the patterns of effects that emerged were generally interpretable, it will be important for findings from these analyses to be replicated in future studies. As well, it remains for future work to see whether our findings will substantively generalize to other cultural or relational contexts, such as in samples of more recently formed romantic couples and romantic couples who are experiencing more pronounced relationship distress.

Ambulatory acoustic recording—specifically in the EAR for Android form factor—has much to recommend it. The hardware is reasonably affordable, familiar to most participants, compact, and does not necessitate the use of a lavalier mic. The app itself is purpose-built, and the Android OS is highly flexible. As previously noted, we did encounter some challenges related to device (firmware) compatibility, which resulted in unexpected and difficult to diagnose and resolve technological failures early in data collection. Still, we were able to collect hundreds of hours of acoustic data. Of course, the subsequent coding is labor intensive. A rough estimate is that coding required over 1500 collective hours of work, not inclusive of trainings and regular coding meetings. Once collected, though, these data potentially lend themselves to multiple forms of analysis, including not only content coding of the type described here, but also acoustic (e.g., vocal prosody) and text analyses, among others. There are, of course, limitations to these data. Participants are aware that they are being audio recorded and may, in some cases, alter their behavior accordingly, potentially including removing the device specifically when they want to have a private conversation with their partner. In addition, ambulatory acoustic recording does not permit direct access to information about facial expressions or body language, including physical touch, which has been shown to play an important role in couples' communication and emotional and physiological coregulation (e.g., Coan et al., 2017; Debrot, Schoebi, Perrez, & Horn, 2013).

In the current investigation, we prioritized ecological validity over internal validity. No experimental manipulations or controls were employed in the current study. Unlike studies using laboratory conflict conversation tasks (e.g., Graber et al., 2011), for example, we did not know whether, when, or where the couples in our study might have a significant disagreement, and we could not control any of the qualities of the acoustic or social environment. Ultimately, I see these approaches as essentially complementary, with important insights emerging from cases in which laboratory studies and studies of daily life yield divergent results, as well as when they yield converging results.

A lingering question is the degree to which observed associations are specific to IER interactions. In the current study, for example, scores on a composite of baseline measures of relationship satisfaction and quality predicted more empathic responding, less harshness and defensiveness, and greater satisfaction in the context of providing IER to one's partner. It is

entirely plausible that these associations would generalize beyond this context to less affectively laden interactions. This question can be addressed in future analyses.

Of note, most of the IER-relevant interactions that we captured were quite brief in duration. A strength of this approach, in other words, is that it allowed to us to capture the kinds of micro-interactions that make up much of couples' day-to-day communication. According to Relational Regulation Theory, for example, "ordinary yet affectively consequential social interactions" are proposed to play a more important role in explaining associations between perceived social support and mental health than interactions that are explicitly geared toward talking about and coping with major stressors (Lakey & Orehek, 2011, p. 6). Such interactions, though, might be liable to go unrecalled or unreported in the context of, for example, a nightly diary study of emotion sharing. To wit, as I have already noted, we frequently observed IER interactions during epochs where both partners reported neither having provided nor received IER. As an empirical matter, it is unclear whether the findings of the current study (e.g., regarding correlations between provider behaviors) would generalize to, for example, couples' interactions in the immediate wake of a major life event. Crucially, though, ratings of these seemingly quotidian interactions were strongly associated with a baseline measure of intimate partner violence, which strongly underscores their relevance.

Conclusion

Notwithstanding limitations, the present investigation makes several notable contributions to the emotion regulation and close relationship literatures, including phenomenological, conceptual, and methodological contributions. By using ambulatory acoustic recording, we gained insight into rich, contextually-embedded affect-laden interactions, many of which were apparently invisible to the participants themselves in that they frequently went unreported in EMA surveys. Among other findings, these data provide compelling evidence that daily interactions between romantic partners are rife with opportunities for IER; that people enact a wide range of behaviors across IER interactions; and that those behaviors vary as a function of contextual factors, but can still be can be predicted from baseline measures, particularly baseline measures that capture relationship dynamics. In the latter regard, the prospective correlations between baseline intimate partner violence and provider behavior codes are particularly striking given the magnitude of these correlations and the devastating costs of intimate partner violence. I hope that the present study will help to create a platform from which future studies can investigate IER in daily life, such as in clinical samples. In conclusion, I have shown that ambulatory acoustic recording is a feasible method for observing IER interactions in daily life and have developed a coding manual that can be used to reliably code the content of these interactions.

References

Abe, K., & Nakashima, K. (2020). Excessive-reassurance seeking and mental health: Interpersonal networks for emotion regulation. *Current Psychology*, 1-11. https://doi.org/10.1007/s12144-020-00955-2

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*, 217-237. https://doi.org/10.1016/j.cpr.2009.11.004

Baddeley, J. L., Pennebaker, J. W., & Beevers, C. G. (2013). Everyday social behavior during a major depressive episode. *Social Psychological and Personality Science*, *4*, 445-452. https://doi.org/10.1177/1948550612461654

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*, 456. https://doi.org/10.3389/fpsyg.2017.00456

Barkham, M., Hardy, G. E., & Startup, M. (1996). The IIP-32: A short version of the Inventory of Interpersonal Problems. *British Journal of Clinical Psychology*, *35*, 21-35. https://doi.org/10.1111/j.2044-8260.1996.tb01159.x

Barrett, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of experienced emotion. *Journal of Personality and Social Psychology*, *87*(2), 266. https://doi.org/10.1037/0022-3514.87.2.266

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Beckes, L., & Coan, J. A. (2012). Social baseline theory and the social regulation of emotion. In L. Campbell, J. G. La Guardia, J. M. Olson, & M. P. Zanna (Eds.), *The science of the couple: The Ontario symposium* (pp. 81-93). New York, NY: Psychology Press. https://doi.org/10.4324/9780203134160-10

Bender, A. K. (2017). Ethics, methods, and measures in intimate partner violence research: the current state of the field. *Violence Against Women*, *23*(11), 1382-1413. https://doi.org/10.1177/1077801216658977

Bloch, L., Haase, C. M., & Levenson, R. W. (2014). Emotion regulation predicts marital satisfaction: More than a wives' tale. *Emotion*, *14*(1), 130. https://doi.org/10.1037/a0034272

Bodenmann, G. (1995). A systemic-transactional conceptualization of stress and coping in couples. *Swiss Journal of Psychology / Schweizerische Zeitschrift für Psychologie / Revue Suisse de Psychologie*, *54*, 34-49.

Bodenmann, G. (2008). *Dyadisches Coping Inventar (DCI): Manual*.

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

Bradbury, T. N., & Fincham, F. D. (1992). Attributions and behavior in marital interaction. *Journal of Personality and Social Psychology*, *63*, 613. https://doi.org/10.1037/0022-3514.63.4.613

Bradbury, T. N., Beach, S. R. H., Fincham, F. D., & Nelson, G. M. (1996). Attributions and behavior in functional and dysfunctional marriages. *Journal of Consulting and Clinical Psychology*, *64*, 569–576. https://doi.org/10.1037/0022-006X.64.3.569

Bratec, S. M., Bertram, T., Starke, G., Brandl, F., Xie, X., & Sorg, C. (2020). Your presence soothes me: a neural process model of aversive emotion regulation via social buffering.

Social Cognitive and Affective Neuroscience, *15*(5), 561-570. https://doi.org/10.1093/scan/nsaa068

Burleson, B. R. (2008). What counts as effective emotional support? Explorations of individual and situational differences. *Studies in Applied Interpersonal Communication*, 207-227. https://doi.org/10.4135/9781412990301.n10

Butler, E. A. (2011). Temporal interpersonal emotion systems: The "TIES" that form relationships. *Personality and Social Psychology Review*, *15*, 367-393. https://doi.org/10.1177/1088868311411164

Butler, E. A. (2015). Interpersonal affect dynamics: It takes two (and time) to tango. *Emotion Review*, *7*, 336-341. https://doi.org/10.1177/1754073915590622

Butler, E. A., Hollenstein, T., Shoham, V., & Rohrbaugh, M. J. (2014). A dynamic state-space analysis of interpersonal emotion regulation in couples who smoke. *Journal of Social and Personal Relationships*, *31*, 907-927. https://doi.org/10.1177/0265407513508732

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193-213. https://doi.org/10.1016/0165-1781(89)90047-4

Caldwell, J. E., Swan, S. C., & Woodbrown, V. D. (2012). Gender differences in intimate partner violence outcomes. *Psychology of Violence*, *2*(1), 42. https://doi.org/10.1037/a0026296

Campbell, J. C. (2002). Health consequences of intimate partner violence. *The Lancet*, *359*(9314), 1331-1336. https://doi.org/10.1016/S0140-6736(02)08336-8

Campos, J. J., Walle, E. A., Dahl, A., & Main, A. (2011). Reconceptualizing emotion regulation. *Emotion Review*, *3*, 26-35. https://doi.org/10.1177/1754073910380975

Carver, C. S., Johnson, S. L., Joormann, J., Kim, Y., & Nam, J. Y. (2011). Serotonin transporter polymorphism interacts with childhood adversity to predict aspects of impulsivity. *Psychological Science*, *22*(5), 589-595. https://doi.org/10.1177/0956797611404085

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., ... & Rose, M. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*, *45*, S3. https://doi.org/10.1097/01.mlr.0000258615.42478.55

Cheung, E. O., Gardner, W. L., & Anderson, J. F. (2015). Emotionships: Examining people's emotion-regulation relationships and their consequences for well-being. *Social Psychological and Personality Science*, *6*, 407-414. https://doi.org/10.1177/1948550614564223

Christensen, K. A., van Dyk, I. S., Nelson, S. V., & Vasey, M. W. (2020). Using multilevel modeling to characterize interpersonal emotion regulation strategies and psychopathology in female friends. *Personality and Individual Differences*, *165*, 110156. https://doi.org/10.1016/j.paid.2020.110156

Christensen, R. H. B. (2019). *ordinal: Regression Models for Ordinal Data (Version 2019.12-10)*. https://CRAN.R-project.org/package=ordinal.

Cloitre, M., Khan, C., Mackintosh, M.-A., Garvert, D. W., Henn-Haase, C. M., Falvey, E. C., & Saito, J. (2019). Emotion regulation mediates the relationship between ACES and physical and mental health. *Psychological Trauma: Theory, Research, Practice, and Policy, 11*(1), 82–89. https://doi.org/10.1037/tra0000374

Coan, J. A., & Gottman, J. M. (2007). The Specific Affect Coding System (SPAFF). In J. Coan & J. Allen (Eds.), Handbook of emotion elicitation and assessment (pp. 267–285). New York: Oxford University Press.

Coan, J. A., Beckes, L., Gonzalez, M. Z., Maresh, E. L., Brown, C. L., & Hasselmo, K. (2017). Relationship status and perceived support in the social regulation of neural responses to threat. *Social Cognitive and Affective Neuroscience, 12*(10), 1574-1583. https://doi.org/10.1093/scan/nsx091

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 385-396. https://doi.org/10.2307/2136404

Compas, B. E., Jaser, S. S., Bettis, A. H., Watson, K. H., Gruhn, M. A., Dunbar, J. P., ... & Thigpen, J. C. (2017). Coping, emotion regulation, and psychopathology in childhood and adolescence: A meta-analysis and narrative review. *Psychological Bulletin*, *143*(9), 939. https://doi.org/10.1037/bul0000110

Dattilio, F. M., & Padesky, C. A. (1990). *Cognitive therapy with couples*. Professional Resource Exchange, Inc.

Debrot, A., Schoebi, D., Perrez, M., & Horn, A. B. (2013). Touch as an interpersonal emotion regulation process in couples' daily lives: the mediating role of psychological intimacy. *Personality and Social Psychology Bulletin, 39*, 1373-1385. https://doi.org/10.1177/0146167213497592

Demiray, B., Luo, M., & Grilli, M. (2020). Sounds of healthy aging: Assessing everyday cognitive activity from real-life audio data. *Innovation in Aging*, *4*, 622-622. https://doi.org/10.1093/geroni/igaa057.2119

DeSteno, D., Gross, J. J., & Kubzansky, L. (2013). Affective science and health: The importance of emotion and emotion regulation. *Health Psychology*, *32*, 474. https://doi.org/10.1037/a0030259

Dixon-Gordon, K. L., Bernecker, S. L., & Christensen, K. (2015). Recent innovations in the field of interpersonal emotion regulation. *Current Opinion in Psychology*, *3*, 36-42. https://doi.org/10.1016/j.copsyc.2015.02.001

Dixon-Gordon, K. L., Haliczer, L. Conkey. L, & Whalen, D. (2018). Difficulties in interpersonal emotion regulation: Initial development and validation of a self-report measure. *Journal of Psychopathology and Behavioral Assessment, 40*, 528-549. https://doi.org/10.1007/s10862-018-9647-9

Dixon-Gordon, K. L., Whalen, D. J., Scott, L. N., Cummins, N. D., & Stepp, S. D. (2016). The main and interactive effects of maternal interpersonal emotion regulation and negative affect on adolescent girls' borderline personality disorder symptoms. *Cognitive Therapy and Research, 40*, 381-393. https://doi.org/10.1007/s10608-015-9706-4

Doré, B. P., Morris, R. R., Burr, D. A., Picard, R. W., & Ochsner, K. N. (2017). Helping others regulate emotion predicts increased regulation of one's own emotions and decreased symptoms of depression. *Personality and Social Psychology Bulletin, 43*, 729–739. https://doi.org/10.1177/0146167217695558

Fearey, E., Evans, J., & Schwartz-Mette, R. A. (2021). Emotion regulation deficits and depression-related maladaptive interpersonal behaviours. *Cognition and Emotion*, *35*(8), 1559-1572. https://doi.org/10.1080/02699931.2021.1989668

Fletcher, G. J., & Kerr, P. S. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin, 136*, 627. https://doi.org/10.1037/a0019792

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. https://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. https://doi.org/10.1037/0893-3200.21.4.572

Gable, S. L., Reis, H. T., Impett, E. A., & Asher, E. R. What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *Journal of Personality and Social Psychology, 87*(2), 228-245. https://doi.org/10.1037/0022-3514.87.2.228

Gamer, M., Lemon, J., Fellows, I., & Puspendra, S. (2012) *irr: Various Coefficients of Interrater Reliability and Agreement (Version 0.84.1)*. https://CRAN.R-project.org/package=irr

Gotlib, I. H. (1983). Perception and recall of interpersonal feedback: Negative bias in depression. *Cognitive Therapy and Research, 7*, 399-412. https://doi.org/10.1007/BF01187168

Gottman, J. M., & Gottman, J. S. (2008). Gottman Method Therapy. In A. S. Gurman (Ed.), *Clinical handbook of couple therapy* (pp. 138-164). New York, NY: Guildford Press.

Gottman, J. M., & Levenson, R. W. (1999). How stable is marital interaction over time?. *Family Process, 38*, 159-165. https://doi.org/10.1111/j.1545-5300.1999.00159.x

Gottman, J. M., & Levenson, R. W. (2000). The timing of divorce: Predicting when a couple will divorce over a 14-year period. *Journal of Marriage and Family, 62*, 737-745. https://doi.org/10.1111/j.1741-3737.2000.00737.x

Gottman, J. M., Coan, J., Carrere, S., & Swanson, C. (1998). Predicting marital happiness and stability from newlywed interactions. *Journal of Marriage and the Family, 60*, 5-22. https://doi.org/10.2307/353438

Graber, E. C., Laurenceau, J. P., Miga, E., Chango, J., & Coan, J. (2011). Conflict and love: predicting newlywed marital outcomes from two interaction contexts. *Journal of Family Psychology, 25*(4), 541. https://doi.org/10.1037/a0024507

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*, 41-54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271-299. https://doi.org/10.1037/1089-2680.2.3.271

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*, 281-291. https://doi.org/10.1017/S0048577201393198

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion, 13*, 359-365. https://doi.org/10.1037/a0032135

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*, 1-26. https://doi.org/10.1080/1047840X.2014.940781

Gross, J. J., & John, O. P. (1997). Revealing feelings: facets of emotional expressivity in self-reports, peer ratings, and behavior. *Journal of Personality and Social Psychology, 72*, 435. https://doi.org/10.1037/0022-3514.72.2.435

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348. https://doi.org/10.1037/0022-3514.85.2.348

Gross, J. J., Richards, J. M., & John, O. P. (2006). Emotion regulation in everyday life. In D. K. Snyder, J. Simpson, & J. N. Hughes (Eds.), *Emotion regulation in couples and families:*

Pathways to dysfunction and health (pp. 13-35). Washington, DC, US: American Psychological Association. https://doi.org/10.1037/11468-001

Haase, C. M., Holley, S. R., Bloch, L., Verstaen, A., & Levenson, R. W. (2016). Interpersonal emotional behaviors and physical health: A 20-year longitudinal study of long-term married couples. *Emotion, 16*, 965. https://doi.org/10.1037/a0040239

Halekoh, U., & Højsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models: The R package pbkrtest. *Journal of Statistical Software, 59*(9), 1–30. https://doi.org/10.18637/jss.v059.i09

Hall, M., Scherner, P. V., Kreidel, Y., & Rubel, J. A. (2021). A systematic review of momentary assessment designs for mood and anxiety symptoms. *Frontiers in Psychology, 12*, 1716. https://doi.org/10.3389/fpsyg.2021.642044

Hallam, G. P., Webb, T. L., Sheeran, P., Miles, E., Niven, K., Wilkinson, I. D., ... & Farrow, T. F. (2014). The neural correlates of regulating another person's emotions: an exploratory fMRI study. *Frontiers in Human Neuroscience, 8*, 376. https://doi.org/10.3389/fnhum.2014.00376

Halperin, E. (2014). Emotion, emotion regulation, and conflict resolution. *Emotion Review, 6*, 68-76. https://doi.org/10.1177/1754073913491844

Hart, B., Risley, T. R., & Kirby, J. R. (1997). Meaningful differences in the everyday experience of young American children. *Canadian Journal of Education, 22*, 323. https://doi.org/10.2307/1585834

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77-89. https://doi.org/10.1080/19312450709336664

Heiy, J. E., & Cheavens, J. S. (2014). Back to basics: A naturalistic assessment of the experience and regulation of emotion. *Emotion, 14*, 878. https://doi.org/10.1037/a0037231

Herbison, J. D., Martin, L. J., Benson, A. J., McLaren, C. D., Slatcher, R. B., Boardley, I. D., ... & Bruner, M. W. (2021). "Beyond the rink": A multilevel analysis of social identity behaviors captured using the Electronically Activated Recorder. *Journal of Sport and Exercise Psychology, 43*(6), 441-449. https://doi.org/10.1123/jsep.2020-0343

Hofmann, S. G. (2014). Interpersonal emotion regulation model of mood and anxiety disorders. *Cognitive Therapy and Research, 38*, 483-492. https://doi.org/10.1007/s10608-014-9620-1

Hokanson, J. E., Rubert, M. P., Welker, R. A., Hollander, G. R., & Hedeen, C. (1989). Interpersonal concomitants and antecedents of depression among college students. *Journal of Abnormal Psychology, 98*, 209. https://doi.org/10.1037/0021-843X.98.3.209

Horn, A. B., & Maercker, A. (2016). Intra-and interpersonal emotion regulation and adjustment symptoms in couples: The role of co-brooding and co-reappraisal. *BMC Psychology, 4*, 51. https://doi.org/10.1186/s40359-016-0159-7

Horn, A. B., Samson, A. C., Debrot, A., & Perrez, M. (2018). Positive humor in couples as interpersonal emotion regulation: a dyadic study in everyday life on the mediating role of psychological intimacy. *Journal of Social and Personal Relationships*, 0265407518788197. https://doi.org/10.1177/0265407518788197

Howland, M., & Simpson, J. A. (2010). Getting in under the radar: A dyadic view of invisible support. *Psychological Science, 21*, 1878-1885. https://doi.org/10.1177/0956797610388817

Jakobs, E., Manstead, A. S. R., & Fischer, A. H. (1999). Social motives and emotional feelings as determinants of facial displays: The case of smiling. *Personality and Social Psychology Bulletin*, *25*, 424–435. https://doi.org/10.1177/0146167299025004003

Jonas, S., Khalifeh, H., Bebbington, P. E., McManus, S., Brugha, T., Meltzer, H., & Howard, L. M. (2014). Gender differences in intimate partner violence and psychiatric disorders in England: results from the 2007 adult psychiatric morbidity survey. *Epidemiology and Psychiatric Sciences*, *23*(2), 189-199. https://doi.org/10.1017/S2045796013000292

Julien, D., Markman, H. J., & Lindahl, K. M. (1989). A comparison of a global and a microanalytic coding system: Implications for future trends in studying interactions. *Behavioral Assessment*, *11*, 81-100.

Kaurin, A., King, K. M., & Wright, A. G. (2022, February 10). Studying personality pathology with ecological momentary assessment: Harmonizing theory and method. PsyArXiv. https://doi.org/10.31234/osf.io/chszq

Kiecolt-Glaser, J. K., Bane, C., Glaser, R., & Malarkey, W. B. (2003). Love, marriage, and divorce: Newlyweds' stress hormones foreshadow relationship changes. *Journal of Consulting and Clinical Psychology*, *71*, 176. https://doi.org/10.1037/0022-006X.71.1.176

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352. https://doi.org/10.1037/h0043688

Kraiss, J. T., Peter, M., Moskowitz, J. T., & Bohlmeijer, E. T. (2020). The relationship between emotion regulation and well-being in patients with mental disorders: A meta-analysis. *Comprehensive psychiatry*, *102*, 152189. https://doi.org/10.1016/j.comppsych.2020.152189

Kresnow, M., Smith, S. G., Basile, K. C., & Chen, J. (2022). *The National Intimate Partner and Sexual Violence Survey: 2016/2017 methodology report*. Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. https://www.cdc.gov/violenceprevention/pdf/nisvs/nisvs-methodology-report 508.pdf

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine*, *42*, 1879-1890. https://doi.org/10.1017/S0033291711002674

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Lakey, B. (2010). Social-clinical approaches to social support suggest new strategies for intervention. In J. E. Maddux & J. P. Tangney (Eds.), *Social psychological foundations of clinical psychology* (pp. 177–194). New York, NY: Guilford Press.

Lakey, B., & Orehek, E. (2011). Relational regulation theory: A new approach to explain the link between perceived social support and mental health. *Psychological Review*, *118*, 482. https://doi.org/10.1037/a0023477

Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, *9*(3), 371-375. https://doi.org/10.1177/0146167283093007

Levy-Gigi, E., & Shamay-Tsoory, S. G. (2017). Help me if you can: Evaluating the effectiveness of interpersonal compared to intrapersonal emotion regulation in reducing distress. *Journal of Behavior Therapy and Experimental Psychiatry*, *55*, 33-40. https://doi.org/10.1016/j.jbtep.2016.11.008

Longford, N. T., & Muthén, B. O. (1992). Factor analysis for clustered observations. *Psychometrika*, *57*(4), 581-597. https://doi.org/10.1007/BF02294421

López-Pérez, B., Ambrona, T., & Gummerum, M. (2017). Interpersonal emotion regulation in Asperger's syndrome and borderline personality disorder. *British Journal of Clinical Psychology*, *56*, 103-113. https://doi.org/10.1111/bjc.12124

Lorenzo, J. M., Barry, R. A., & Khalifian, C. E. (2018). More or less: Newlyweds' preferred and received social support, affect, and relationship satisfaction. *Journal of Family Psychology*, *32*, 860. https://doi.org/10.1037/fam0000440

Lüdecke, D. (2021). *sjstats: Statistical Functions for Regression Models (Version 0.18.1)*. http://doi.org/10.5281/zenodo.1284472, https://CRAN.R-project.org/package=sjstats

Maisel, N. C., & Gable, S. L. (2009). The paradox of received social support: The importance of responsiveness. *Psychological Science*, *20*, 928-932. https://doi.org/10.1111/j.1467-9280.2009.02388.x

Manson, J. H., & Robbins, M. L. (2017). New evaluation of the electronically activated recorder (EAR): obtrusiveness, compliance, and participant self-selection effects. *Frontiers in Psychology*, *8*, 658. https://doi.org/10.3389/fpsyg.2017.00658

Margolin, G., & Weiss, R. L. (1978). Comparative evaluation of therapeutic components associated with behavioral marital treatments. *Journal of Consulting and Clinical Psychology*, *46*, 1476. https://doi.org/10.1037/0022-006X.46.6.1476

Marroquín, B. (2011). Interpersonal emotion regulation as a mechanism of social support in depression. *Clinical Psychology Review*, *31*, 1276-1290. https://doi.org/10.1016/j.cpr.2011.09.005

Martínez-Íñigo, D., Poerio, G. L., & Totterdell, P. (2013). The association between controlled interpersonal affect regulation and resource depletion. *Applied Psychology: Health and Well-Being*, *5*, 248-269. https://doi.org/10.1111/aphw.12009

Matos, M., Pinto-Gouveia, J., Gilbert, P., Duarte, C., & Figueiredo, C. (2015). The Other As Shamer Scale–2: Development and validation of a short version of a measure of external shame. *Personality and Individual Differences*, *74*, 6-11. https://doi.org/10.1016/j.paid.2014.09.037

McNeish, D. (2017). Small sample methods for multilevel modeling: A colloquial elucidation of REML and the Kenward-Roger correction. *Multivariate Behavioral Research*, *52*(5), 661-670. https://doi.org/10.1080/00273171.2017.1344538

Mehl, M. R. (2017). The electronically activated recorder (EAR) a method for the naturalistic observation of daily social behavior. *Current Directions in Psychological Science*, *26*(2), 184-190. https://doi.org/10.1177/0963721416680611

Mehl, M. R., & Holleran, S. E. (2007). An empirical analysis of the obtrusiveness of and participants' compliance with the electronically activated recorder (EAR). *European Journal of Psychological Assessment*, *23*, 248-257. https://doi.org/10.1027/1015-5759.23.4.248

Mehl, M. R., & Robbins, M. L. (2012). Naturalistic observation sampling: The Electronically Activated Recorder (EAR). In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 176-192). New York: Guilford Press.

Mellor, R. M., Slaymaker, E., & Cleland, J. (2013). Recognizing and overcoming challenges of couple interview research. *Qualitative Health Research*, *23*(10), 1399-1407. https://doi.org/10.1177/1049732313506963

Minor, K. S., Davis, B. J., Marggraf, M. P., Luther, L., & Robbins, M. L. (2018). Words matter: Implementing the electronically activated recorder in schizotypy. *Personality Disorders: Theory, Research, and Treatment*, *9*(2), 133. https://doi.org/10.1037/per0000266

Miu, A. C., Szentágotai-Tătar, A., Balazsi, R., Nechita, D., Bunea, I., & Pollak, S. D. (2022). Emotion regulation as mediator between childhood adversity and psychopathology: A meta-analysis. *Clinical Psychology Review*, 102141. https://doi.org/10.1016/j.cpr.2022.102141

Morris, A. S., Silk, J. S., Morris, M. D., Steinberg, L., Aucoin, K. J., & Keyes, A. W. (2011). The influence of mother–child emotion regulation strategies on children's expression of anger and sadness. *Developmental Psychology*, *47*, 213. https://doi.org/10.1037/a0021021

Murray, S. L., & Holmes, J. G. (1997). A leap of faith? Positive illusions in romantic relationships. *Personality and Social Psychology Bulletin*, *23*, 586-604. https://doi.org/10.1177/0146167297236003

Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996a). The benefits of positive illusions: Idealization and the construction of satisfaction in close relationships. *Journal of Personality and Social Psychology*, *70*, 79. https://doi.org/10.1037/0022-3514.70.1.79

Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996b). The self-fulfilling nature of positive illusions in romantic relationships: Love is not blind, but prescient. *Journal of Personality and Social Psychology*, *71*, 1155. https://doi.org/10.1037/0022-3514.71.6.1155

Nelson, G. H., O'Hara, M. W., & Watson, D. (2018). National norms for the expanded version of the inventory of depression and anxiety symptoms (IDAS-II). *Journal of Clinical Psychology*, *74*(6), 953-968. https://doi.org/10.1002/jclp.22560

Netzer, L., Van Kleef, G. A., & Tamir, M. (2015). Interpersonal instrumental emotion regulation. *Journal of Experimental Social Psychology*, *58*, 124-135. https://doi.org/10.1016/j.jesp.2015.01.006

Nils, F., & Rimé, B. (2012). Beyond the myth of venting: Social sharing modes determine the benefits of emotional disclosure. *European Journal of Social Psychology*, *42*, 672-681. https://doi.org/10.1002/ejsp.1880

Niven, K. (2017). The four key characteristics of interpersonal emotion regulation. *Current Opinion in Psychology*, *17*, 89-93. https://doi.org/10.1016/j.copsyc.2017.06.015

Niven, K., Garcia, D., van der Löwe, I., Holman, D., & Mansell, W. (2015). Becoming popular: interpersonal emotion regulation predicts relationship formation in real life social networks. *Frontiers in Psychology*, *6*, 1452. https://doi.org/10.3389/fpsyg.2015.01452

Niven, K., Holman, D., & Totterdell, P. (2012). How to win friendship and trust by influencing people's feelings: An investigation of interpersonal affect regulation and the quality of relationships. *Human Relations*, *65*, 777-805. https://doi.org/10.1177/0018726712439909

Niven, K., Macdonald, I., & Holman, D. (2012). You spin me right round: Cross-relationship variability in interpersonal emotion regulation. *Frontiers in Psychology*, *3*, 394. https://doi.org/10.3389/fpsyg.2012.00394

Niven, K., Totterdell, P., & Holman, D. (2009). A classification of controlled interpersonal affect regulation strategies. *Emotion*, *9*, 498. https://doi.org/10.1037/a0015962

Niven, K., Totterdell, P., Holman, D., & Headley, T. (2012). Does regulating others' feelings influence people's own affective well-being? *The Journal of Social Psychology*, *152*, 246-260. https://doi.org/10.1080/00224545.2011.599823

Niven, K., Troth, A. C., & Holman, D. (2019). Do the effects of interpersonal emotion regulation depend on people's underlying motives?. *Journal of Occupational and Organizational Psychology, 92*(4), 1020-1026. https://doi.org/10.1111/joop.12257

O'Brien, T. B., DeLongis, A., Pomaki, G., Puterman, E., & Zwicker, A. (2009). Couples coping with stress: The role of empathic responding. *European Psychologist, 14*(1), 18-28. https://doi.org/10.1027/1016-9040.14.1.18

Operario, D., Adler, N. E., & Williams, D. R. (2004). Subjective social status: Reliability and predictive utility for global health. *Psychology & Health, 19*(2), 237-246. https://doi.org/10.1080/08870440310001638098

Parlamis, J. D. (2012). Venting as emotion regulation: The influence of venting responses and respondent identity on anger and emotional tone. *International Journal of Conflict Management, 23*(1), 77–96. https://doi.org/10.1108/10444061211199322

Pauw, L. S., Sauter, D. A., van Kleef, G. A., & Fischer, A. H. (2017). Sense or sensibility? Social sharers' evaluations of socio-affective vs. cognitive support in response to negative emotions. *Cognitive and Emotion, 32*, 1247-1264. https://doi.org/10.1080/02699931.2017.1400949

Pauw, L. S., Sauter, D. A., van Kleef, G. A., & Fischer, A. H. (2019). Stop crying! The impact of situational demands on interpersonal emotion regulation. *Cognition and Emotion, 33*(8), 1587-1598. https://doi.org/10.1080/02699931.2019.1585330

Reeck, C., Ames, D. R., & Ochsner, K. N. (2016). The social regulation of emotion: An integrative, cross-disciplinary model. *Trends in Cognitive Sciences, 20*, 47-63. https://doi.org/10.1016/j.tics.2015.09.003

Reis, H. T., & Gable, S. L. (2015). Responsiveness. *Current Opinion in Psychology, 1*, 67-71. https://doi.org/10.1016/j.copsyc.2015.01.001

Reis, H. T., Crasta, D., Rogge, R. D., Maniaci, M. R., & Carmichael, C. L. (in press). Perceived partner responsiveness scale. In D. L. Worthington & G. D. Bodie (Eds.), *The sourcebook of listening research methodology and measures*. New York: John Wiley & Sons, Inc.

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review, 1*, 60-85. https://doi.org/10.1177/1754073908097189

Robbins, M. L., & Karan, A. (2020). Who gossips and how in everyday life?. *Social Psychological and Personality Science, 11*(2), 185-195. https://doi.org/10.1177/1948550619837000

Robbins, M. L., López, A. M., Weihs, K. L., & Mehl, M. R. (2014). Cancer conversations in context: Naturalistic observation of couples coping with breast cancer. *Journal of Family Psychology, 28*, 380. https://doi.org/10.1037/a0036458

Robbins, M. L., Mehl, M. R., Holleran, S. E., & Kasle, S. (2011). Naturalistically observed sighing and depression in rheumatoid arthritis patients: a preliminary study. *Health Psychology, 30*(1), 129. https://doi.org/10.1037/a0021558

Robbins, M. L., Wright, R. C., María López, A., & Weihs, K. (2019). Interpersonal positive reframing in the daily lives of couples coping with breast cancer. *Journal of Psychosocial Oncology, 37*(2), 160-177. https://doi.org/10.1080/07347332.2018.1555198

Robbins, Wright, Karan, & Baranski (2021). *EAR repository*. Open Science Foundation. https://osf.io/n2ufd/

Rose, A. J., Carlson, W., & Waller, E. M. (2007). Prospective associations of co-rumination with friendship and emotional adjustment: considering the socioemotional trade-offs of co-

rumination. *Developmental Psychology*, *43*, 1019. https://doi.org/10.1037/0012-1649.43.4.1019

Ruan, Y., Reis, H. T., Clark, M. S., Hirsch, J. L., & Bink, B. D. (2020). Can I tell you how I feel? Perceived partner responsiveness encourages emotional expression. *Emotion*, *20*(3), 329. https://doi.org/10.1037/emo0000650

Rubin, Z. (1970). Measurement of romantic love. *Journal of Personality and Social Psychology*, *16*, 265. https://doi.org/10.1037/h0029841

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161. https://doi.org/10.1037/h0077714

Sahi, R. S., Ninova, E., & Silvers, J. A. (2020). With a little help from my friends: Selective social potentiation of emotion regulation. *Journal of Experimental Psychology: General*, *150*(6), 1237-1249. https://doi.org/10.1037/xge0000853

Sanford, K. (2006). Communication during marital conflict: When couples alter their appraisal, they change their behavior. *Journal of Family Psychology*, *20*, 256-265. https://doi.org/10.1037/0893-3200.20.2.256

Schachter, S. (1959). *The psychology of affiliation: Experimental studies of the sources of gregariousness*. Stanford, CA: Stanford University Press

Schwarzer, R., & Jerusalem, M. (1995). Optimistic self-beliefs as a resource factor in coping with stress. In S. E. Hobfoll & M. W. de Vries (Eds.), *Extreme stress and communities: Impact and intervention* (pp. 159-177). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-015-8486-9_7

Shaver, P. R., & Mikulincer, M. (2007). Adult attachment strategies and the regulation of emotion. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 446-465). New York, NY: Guilford Press.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1-32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shu, J., Bolger, N., & Ochsner, K. N. (2021). Social emotion regulation strategies are differentially helpful for anxiety and sadness. *Emotion, 21*(6), 1144–1159. https://doi.org/10.1037/emo0000921

Slatcher, R. B., & Robles, T. F. (2012). Preschoolers' everyday conflict at home and diurnal cortisol patterns. *Health Psychology*, *31*(6), 834. https://doi.org/10.1037/a0026774

Smith, S.G., Zhang, X., Basile, K.C., Merrick, M.T., Wang, J., Kresnow, M., Chen, J. (2018). *The National Intimate Partner and Sexual Violence Survey (NISVS): 2015 data brief – updated release*. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention. https://www.cdc.gov/violenceprevention/pdf/2015data-brief508.pdf

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*, 117-143. https://doi.org/10.1037/pspp0000096

Spendelow, J. S., Simonds, L. M., & Avery, R. E. (2017). The relationship between co-rumination and internalizing problems: A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, *24*(2), 512-527. https://doi.org/10.1002/cpp.2023

Stasik-O'Brien, S. M., Brock, R. L., Chmielewski, M., Naragon-Gainey, K., Koffel, E., McDade-Montez, E., ... & Watson, D. (2019). Clinical utility of the inventory of

depression and anxiety symptoms (IDAS). *Assessment*, *26*(5), 944-960. https://doi.org/10.1177/1073191118790036

Stone, L. B., Hankin, B. L., Gibb, B. E., & Abela, J. R. (2011). Co-rumination predicts the onset of depressive disorders during adolescence. *Journal of Abnormal Psychology*, *120*, 752. https://doi.org/10.1037/a0023384

Swerdlow, B. A., & Johnson, S. L. (2020). The Interpersonal Regulation Interaction Scale (IRIS): A multistudy investigation of receivers' retrospective evaluations of interpersonal emotion regulation interactions. *Emotion*. https://doi.org/10.1037/emo0000927

Swerdlow, B. A., & Johnson, S. L. (2016). *Wanting more, but finding less: Associations of psychological symptoms with interpersonal emotion regulation* [Poster presentation]. Society for Research in Psychopathology, Baltimore, MD.

Swerdlow, B. A., & Johnson, S. L. (2021). Associations between intra- and inter-personal emotion regulation. *Affective Science*. https://doi.org/10.1007/s42761-021-00071-9

Swerdlow, B. A. & Johnson, S. L. (in preparation). Interpersonal emotion regulation in daily life: Associations with perceived benefits, psychological well-being, social connectedness, and future support-seeking.

Swerdlow, B. A., Berk, L., Johnson, S. L. (in preparation). Experiences of interpersonal emotion regulation: A cross-sample comparison of college students, adults with emotionally impulsive aggression, and adults with lifetime bipolar diagnoses.

Swerdlow, B. A., Sandel, D., & Johnson, S. L. (in press). Shame on me for needing you: Links between interpersonal emotion regulation and experiences of shame. *Emotion*.

Tamir, M., John, O. P., Srivastava, S., & Gross, J. J. (2007). Implicit theories of emotion: Affective and social outcomes across a major life transition. *Journal of Personality and Social Psychology*, *92*, 731. https://doi.org/10.1037/0022-3514.92.4.731

Tanner, S. M., Lakey, B., Cohen, J. L., MacGeorge, E. L., Clark, R. A., Stewart, S. & Robinson, L. (2018). What is the right thing to say?: Agreement among perceivers on the supportiveness of statements. *Basic and Applied Social Psychology*, *40*, 329-339. https://doi.org/10.1080/01973533.2018.1509341

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the Society for Research in Child Development*, *59*, 25-52. https://doi.org/10.1111/j.1540-5834.1994.tb01276.x

Tomko, R. L., Brown, W. C., Tragesser, S. L., Wood, P. K., Mehl, M. R., & Trull, T. J. (2014). Social context of anger in borderline personality disorder and depressive disorders: Findings from a naturalistic observation study. *Journal of Personality Disorders*, *28*, 434-448. https://doi.org/10.1521/pedi_2012_26_064

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, *90*, 288. https://doi.org/10.1037/0022-3514.90.2.288

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth edition). Springer, New York. https://doi.org/10.1007/978-0-387-21706-2

Watson, D., O'Hara, M. W., Naragon-Gainey, K., Koffel, E., Chmielewski, M., Kotov, R., ... & Ruggero, C. J. (2012). Development and validation of new anxiety and bipolar symptom scales for an expanded version of the IDAS (the IDAS-II). *Assessment*, *19*, 399-420. https://doi.org/10.1177/1073191112449857

Watson, D., O'Hara, M. W., Chmielewski, M., McDade-Montez, E. A., Koffel, E., Naragon, K., & Stuart, S. (2008). Further validation of the IDAS: evidence of convergent,

discriminant, criterion, and incremental validity. *Psychological Assessment*, *20*(3), 248. https://doi.org/10.1037/a0012570

Watson, D., O'Hara, M. W., Simms, L. J., Kotov, R., Chmielewski, M., McDade-Montez, E. A., ... & Stuart, S. (2007). Development and validation of the Inventory of Depression and Anxiety Symptoms (IDAS). *Psychological Assessment*, *19*(3), 253. https://doi.org/10.1037/1040-3590.19.3.253

Webb, T. L., Miles, E., & Sheeran, P. (2012). Dealing with feeling: a meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychological Bulletin*, *138*(4), 775. https://doi.org/10.1037/a0027600

Weitzner, D., Calamia, M., Stasik-O'Brien, S. M., De Vito, A., & Pugh, E. (2020). Psychometric properties of the expanded version of the inventory of depression and anxiety symptoms (IDAS-II) in a sample of older adults. *Aging & Mental Health*, *24*(11), 1847-1853. https://doi.org/10.1080/13607863.2019.1636206

Wells, J. L., Haase, C. M., Rothwell, E. S., Naugle, K. G., Otero, M. C., Brown, C. L., Lai, J., Chen, K.-H., Connelly, D. E., Grimm, K. J., Levenson, R. W., & Fredrickson, B. L. (2022). Positivity resonance in long-term married couples: Multimodal characteristics and consequences for health and longevity. *Journal of Personality and Social Psychology.* Advance online publication. https://doi.org/10.1037/pspi0000385

Wenzel, M., Rowland, Z., Weber, H., & Kubiak, T. (2020). A round peg in a square hole: strategy-situation fit of intra-and interpersonal emotion regulation strategies and controllability. *Cognition and Emotion*, *34*(5), 1003-1009. https://doi.org/10.1080/02699931.2019.1697209

Wester, R. A., Rubel, J., Zimmermann, J., Hall, M., Kaven, L., & Watson, D. (2021, November 15). Development and Validation of the Inventory of Depression and Anxiety Symptoms II – German Version. https://doi.org/10.31234/osf.io/dnjzu

Wiley, J. F. (2020). *multilevelTools: Multilevel and mixed effects model diagnostics and effect sizes (Version 0.1.1). https://CRAN.R-project.org/package=multilevelTools*

Williams, W. C., Morelli, S. A., Ong, D. C., & Zaki, J. (2018). Interpersonal emotion regulation: Implications for affiliation, perceived support, relationships, and well-being. *Journal of Personality and Social Psychology*, *115*, 656. https://doi.org/10.1037/pspi0000132

Zaki, J., & Williams, W. C. (2013). Interpersonal emotion regulation. *Emotion*, *13*, 803. https://doi.org/10.1037/a0033839

Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science*, *19*, 399-404. https://doi.org/10.1111/j.1467-9280.2008.02099.x

Zee, K. S., & Bolger, N. (2019). Visible and invisible social support: How, why, and when. *Current Directions in Psychological Science*, *28*(3), 314-320. https://doi.org/10.1177/0963721419835214

Zee, K. S., Bolger, N., & Higgins, E. T. (2020). Regulatory effectiveness of social support. *Journal of Personality and Social Psychology*, *119*(6), 1316. https://doi.org/10.1037/pspi0000235

Table 1.

*Descriptive Statistics for Key Variables (*N = *116)*

	M	SD	Possible Range	Skewness	Kurtosis	Within-Dyad Correlation
CSI-16 Relationship Satisfaction	63.69	15.68	0-81	-1.36	4.77	.62
RLS Romantic Love	87.16	14.27	13-117	-.42	2.54	.20
PPRS Responsiveness	125.26	30.34	18-162	-1.34	4.79	.58
DCI Evaluations of Dyadic Coping	11.87	2.75	3-15	-0.93	3.82	.61
Intimate Partner Violence	3.07	7.15	0-145	3.96	20.19	–
IDAS-II Appetite Loss	5.54	2.97	3-15	1.10	3.45	.07
IDAS-II Dysphoria	22.07	7.32	10-50	0.58	2.79	-.06
IDAS-II Well-Being	24.54	6.08	8-40	0.07	2.41	.14
IDAS-II Lassitude	14.20	4.74	6-30	0.43	2.78	.25
IDAS-II Panic	12.54	4.72	8-40	1.50	4.80	.04
IDAS-II Insomnia	13.03	6.12	6-30	0.93	3.34	.27
IDAS-II Ill Temper	8.18	3.65	5-25	2.17	8.24	.21
IDAS-II Suicidality	7.69	3.13	6-30	2.52	9.11	.11
IDAS-II Social Anxiety	11.95	4.83	6-30	0.82	3.08	-.06
IDAS-II General Depression	43.57	12.68	20-100	0.88	3.46	.14
PSS-4 Perceived Stress	10.47	2.92	4-20	0.27	2.80	.17
DERS Emotion Dysregulation	83.05	21.79	36-180	0.55	2.54	.03
ERQ Reappraisal	28.04	7.24	6-42	-0.82	3.90	.01
ERQ Suppression	13.45	5.20	4-28	0.15	2.43	-.05

Note. CSI-16 = Couples Satisfaction Index-16; PPRS = Perceived Partner Responsiveness Scale; DCI = Dyadic Coping Inventory; IDAS-II = The Inventory of Depression and Anxiety Symptoms—Second Version; PSS-4 = Perceived Stress Scale-4; DERS = Difficulties in Emotion Regulation Scale; ERQ = Emotion Regulation Questionnaire.

Table 2.

*Bivariate correlations between key baseline variables (*N = *116)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Relationship Quality	–	-.44***	-.36***	-.43***	-.17	-.01	-.28**
(2) Intimate Partner Violence		–	.11	.14	.05	.13	.19
(3) Internalizing Symptoms			–	.67***	.55***	-.12	.26**
(4) PSS-4 Perceived Stress				–	.55***	-.19	.24*
(5) DERS Emotion Dysregulation					–	-.27**	.31**
(6) ERQ Reappraisal Use						–	.10
(7) ERQ Suppression Use							–

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. PSS-4 = Perceived Stress Scale-4; DERS = Difficulties in Emotion Regulation Scale; ERQ = Emotion Regulation Questionnaire

Table 3.

Exploratory bivariate analyses predicting self-reported obtrusiveness of and percent time spent wearing the EAR (*N* = 114**)**

	Obtrusiveness-Self	Obtrusiveness-Other
Key Variables		
Relationship Quality	-.13	-.18
Intimate Partner Violence	-.04	-.30**
Internalizing Symptoms	.09	.11
PSS-4 Perceived Stress	-.08	-.12
DERS Emotion Dysregulation	-.11	-.12
ERQ Reappraisal	-.22*	-.01
ERQ Suppression	.13	-.03
Participant Demographics		
Gender (0 = Male, 1 = Female)	.28**	-.17
Age	.44***	.08
Years of Education	.42***	.11
Subjective Socioeconomic Status	.12	-.08
Relationship Duration	.44***	-.08
Children (0 = None, 1 = 1+)	.19*	-.08
Confounds and Moderators		
BEQ Positive Expressivity	-.34***	-.15
BEQ Negative Expressivity	.01	.11
BFI-2 Extraversion	-.33***	-.10
BFI-2 Negative Affectivity	.03	.04
BFI-2 Agreeableness	.02	-.03
FNES Fear of Negative Evaluation	.21*	.18*
OAS-2 External Shame Proneness	.32***	.17
BAE Can Control Emotions	.01	.01
BAE Should Control Emotions	.19	-.01
GSE General Self-Efficacy	-.26**	-.18*
ECR-R Attachment Insecurity	.10	.31**
IIP-32 Interpersonal Problems	.24*	.19*
Perceived Emotional Support	-.02	.04

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. PSS-4 = Perceived Stress Scale-4; DERS = Difficulties in Emotion Regulation Scale; ERQ = Emotion Regulation Questionnaire; BEQ = Berkeley Expressivity Questionnaire; BFI-2 = Big Five Inventory-2; FNES = Fear of Negative Evaluation Scale; OAS-2 = Other as Shamer-2; BAE = Beliefs About Emotions; GSE = General Self-Efficacy Scale; ECR-R = Experiences in Close Relationships-Revised; IIP-32 = Inventory of Interpersonal Problems-32.

Table 4.

Inter-rater reliability statistics for first- and second-pass EAR codes

	Agreement (%)	Agreement (Krippendorff's α)	Consistency (ICC)
IER-Relevant (First Pass)	87	.84	
Locus of the Interaction	78	.54	
Specific Emotion	79	.70	
Empathic Responding			0.66
Harsh Responding			0.76
Providing Distractions			0.60
Providing Reassurance			0.67
Practical Support			0.92
Cognitive Support			0.64
Expressions of Warmth			0.51
Defensive Responding			0.71
Expressive Suppression			0.66
(Co-)Brooding			0.59
Emotional Intensity (Start)			0.59
Emotional Intensity (End)			0.62
Receiver Satisfaction			0.79
Provider Satisfaction			0.79
Attunement			0.54

Table 5.

*Skewness, kurtosis, and within-person intraclass correlations for second pass provider behavior codes (*k *= 821 interactions)*

	Skewness	Kurtosis	ICC
Empathic Responding	1.20	4.27	.20
Harsh Responding	2.49	9.13	.38
Providing Distractions	7.35	7.39	.20
Providing Reassurance	4.98	3.56	.16
Practical Support	2.35	8.47	.11
Cognitive Support	2.15	7.57	.19
Expressions of Warmth	1.52	5.39	.10
Defensiveness	2.85	11.53	.24
Expressive Suppression	11.55	15.13	.57
(Co)-Brooding	7.87	7.12	.34

Table 6.

*Within-person correlations between provider behavior codes (*k = *821 interactions)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Empathic	–	-.14***	.13***	.03	-.04	.04	.06*	-.11***	-.13***	-.02
(2) Harsh Responding		–	-.01	.02	.01	.08**	-.07	.29***	.19***	.22***
(3) Providing			–	.13	.14***	.08**	.09**	.07**	.03	.04
(4) Providing				–	.02	.11***	.04	.00	.02	.03
(5) Practical Support					–	.04	.08**	.07**	.11***	.10**
(6) Cognitive Support						–	.11***	.10**	.09**	.13
(7) Expressions of							–	.04	.01	.03
(8) Defensiveness								–	.28***	.26***
(9) Expressive									–	.15***
(10) (Co-)Brooding										–

Table 7.

*Within-person associations between coder ratings of provider behaviors, emotional valence, interaction locus, and mutual regulation expressed as odds ratios (*k *= 821 interactions)*

	Emotional Valence	Interaction Locus	Cause	Mutually Regulatory
Empathic Responding	1.35	.26***	.30***	.66
Harsh Responding	5.37***	9.45***	3.22***	4.74***
Providing Distractions	6.66***	.93	.26***	1.41
Providing Reassurance	4.57***	2.12**	.71	1.68
Practical Support	4.02***	1.37	.68	1.15
Cognitive Support	3.46***	1.95**	.67	1.33
Expressions of Warmth	.31***	1.73*	1.20	1.05
Defensiveness	6.11***	16.51***	5.02***	3.96***
Expressive Suppression	8.91***	1.44	1.11	5.12***
(Co)-Brooding	19.96***	1.53	2.60**	2.01*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. Emotional Valence was dummy-coded as 0 = Positive, 1 = Negative. Interaction Locus was dummy-coded as 0 = External, 1 = Internal. Cause was dummy-coded as 0 = Provider was not the cause, 1 = Provider was the cause for the subset of interactions where the locus was internal. Mutually Regulatory was dummy-coded as 0 = Not mutually regulatory, 1 = Mutually regulatory.

Table 8.

*Within-person correlations between coder ratings of provider behaviors, emotional intensity, and satisfaction (*k = *821 interactions)*

	Emotional Intensity (Start)	Emotional Intensity (End)	Emotional Intensity (Δ)	Receiver Satisfaction	Provider Satisfaction
Empathic Responding	-.02	-.08*	.06*	.21***	.22***
Harsh Responding	.18***	.21***	-.03	-.24***	-.33***
Providing Distractions	.04	-.12**	.13**	.18***	.15**
Providing Reassurance	.10*	.04	.06	.00	.02
Practical Support	-.00	.01	-.01	.01	-.02
Cognitive Support	.04	.05	.00	.04	.03
Expressions of Warmth	.00	-.04	.05	.24***	.32***
Defensiveness	.14**	.15**	-.01	-.18***	-.21***
Expressive Suppression	.31***	.24***	.06	-.21***	-.31***
(Co)-Brooding	.15**	.12**	.04	-.15**	-.20***

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

Table 9.

Time-linked correlations between selected EAR codes and EMA ratings (*k = 162 interactions*)

	Provider Self-Report	Receiver Self-Report
Empathic Responding	.06	.13*
Harsh Responding	.17*	.11*
Cognitive Support	.12*	.08
Providing Distractions	.01	.08
Provider Satisfaction	.16*	—
Receiver Satisfaction	—	.17*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. The variables selected for these analyses were those where there was a conceptual match between pairs of EAR and EMA variables. Analyses were time-linked by epoch.

Table 10.

Correlates between within-person aggregates of selected EAR codes and EMA ratings (N = 97)

	Provider Self-Report	Receiver Self-Report
Empathic Responding	.21*	.41***
Harsh Responding	.23*	.24*
Cognitive Support	.14	.23*
Distraction	.22*	.18
Provider Satisfaction	.13	—
Receiver Satisfaction	—	.10

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

Table 11.

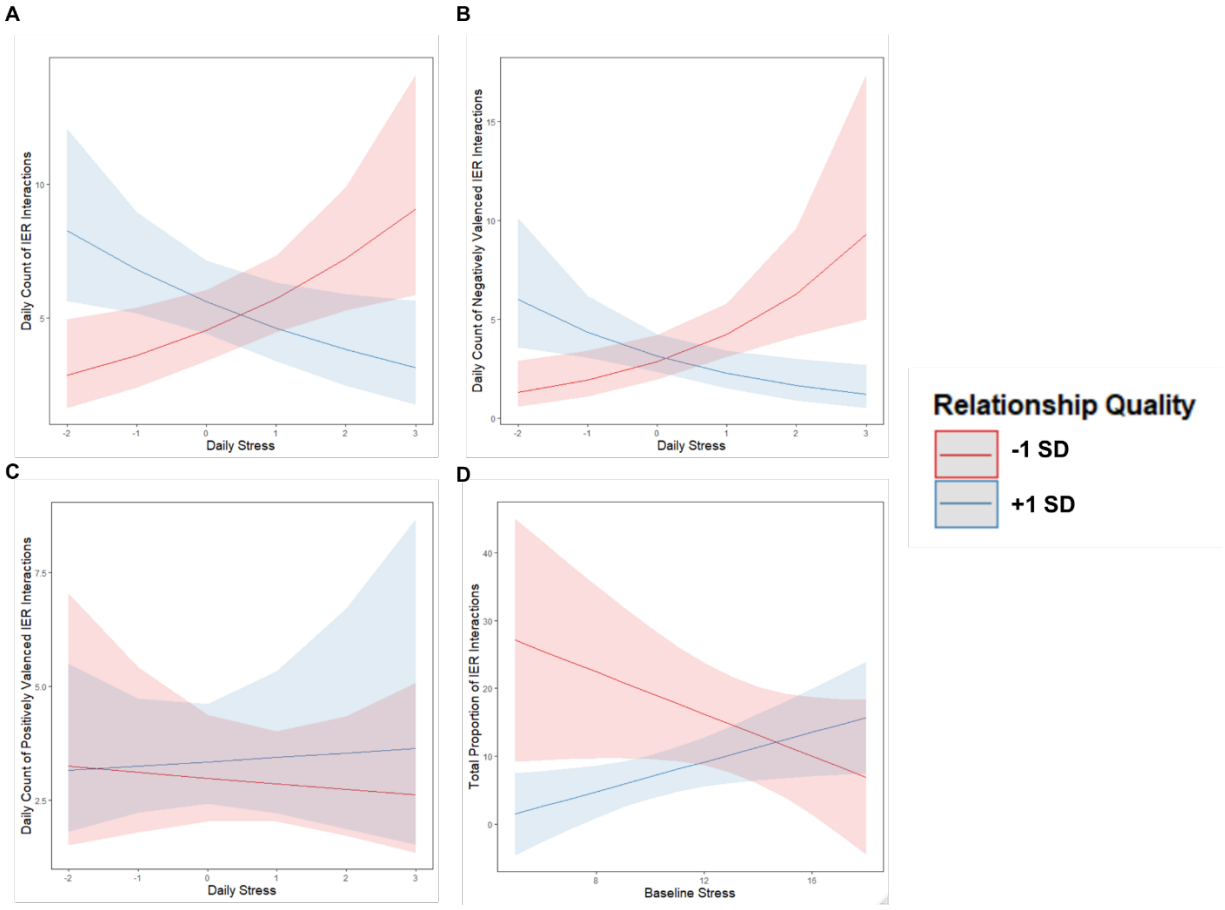
*Correlations between key baseline measures and aggregated EAR codes (*N = *97)*

	Relationship Quality	Intimate Partner Violence	Internalizing Symptoms	PSS-4 Perceived Stress	DERS Emotion Dysregulation	ERQ Reappraisal	ERQ Suppression
Empathic Responding	.25*	-.67**	.09	.07	-.03	-.16	-.10
Harsh Responding	-.38***	.55***	-.10	-.10	-.07	-.04	-.12
Providing Distractions	.08	-.04	.05	.07	.13	-.03	-.06
Providing Reassurance	.11	.16	-.23*	-.20*	-.20*	.07	-.07
Practical Support	-.03	.24	-.13	-.05	.06	-.25*	-.13
Cognitive Support	.18	.14	.12	-.16	.07	-.09	-.13
Expressions of Warmth	.18	-.62***	.10	-.12	-.06	-.04	.03
Defensiveness	-.25*	.48***	-.01	-.08	-.12	-.03	.24*
Expressive Suppression	-.10	-.27	-.10	-.04	-.03	.03	.20*
(Co)-Brooding	-.19	.63***	-.01	-.15	-.17	.29**	-.15
Provider Satisfaction	.26**	-.57***	.10	.11	.00	.02	.02
Receiver Satisfaction	.34***	-.65***	.01	-.01	-.02	-.03	.09

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. PSS-4 = Perceived Stress Scale-4; DERS = Difficulties in Emotion Regulation Scale; ERQ = Emotion Regulation Questionnaire

Figure 1.

Associations between stress, relationship quality, and their interaction and the frequency of IER interactions.



Appendix A: EAR IER Coding Training Manual

Processing and Coding of Audio Files

Note: If at any time you encounter *any* identifying information (e.g., full names other than the participants' names; addresses; phone numbers; etc.), use Audacity to remove the relevant audio segment.

Note: If one of the recordings that you listen to contains overt evidence that a participant intends to end their life imminently, intends to end someone else's life imminently, or is abusing a child, notify [redacted] immediately by phone.

What you'll need:

1. **Audio Processing Logbook (Excel Spreadsheet)**: A logbook to track when and by whom each audio recording is processed. Each row should represent a single audio file. Note the filename and stage of processing (First Pass Completed) with the date at which that stage was completed and the name of the research assistant who completed that stage.
2. **Interaction Record (Excel Spreadsheet)**: A record of each interaction between the couple. Each row should represent a single interaction. Consists of the Couples ID (e.g., if the members of the couple are 101A and 101B, then their Couples ID is 101), the filename, the start time and end time of the interaction. Assign each recording a sequential Recording ID. Note date of First Coding and the research assistant who completed the first coding. Note date of Second Coding and the research assistant who completed the second coding.
3. **Coding Workbook (Qualtrics Form)**: Each submission to the Coding Workbook will reflect a second pass of a single interaction. Consists of the Couples ID, the Recording ID, the coder's name, and the complete list of second codes.
4. **The Audio Recordings**
5. **Audacity** (https://www.audacityteam.org/). Audacity is a free, multiplatform audio editing tool.
6. **7-Zip** (https://www.7-zip.org/). 7-zip is a free, multiplatform tool that will allow you to extract the audio files, which are contained in encrypted zip folders. The password is [redacted].
7. **Headphones**. You **must** wear headphones any time that you are listening to audio recordings for coding purposes.

First Pass:

1. Refer to the **Audio Processing Logbook**. Find the next audio file that has not yet undergone the first pass.
2. Each audio file should be ten minutes in duration. Listen to the first minute of the recording.

i. If there is *no* human conversation taking place in that first minute, you've completed the first pass for this audio file. Make the appropriate notation in the Audio Processing Logbook and move on to the next audio file.

ii. If there is ***any*** human conversation taking place, continue listening to the interaction for long enough to ascertain whether it is an interaction between the couple. **Note:** determining whether the interaction is between the couple may require listening beyond the first minute, listening to the previous or subsequent audio recording, or listening to the relevant section of the paired audio recording (e.g., Partner B's contemporaneous recording) to see if the conversation appears on that audio recording as well. Be careful to make sure you're not hearing a TV show, movie, or similar. We're collecting "baseline audio samples" for each participant so that you will know what their voices sound like; refer to these samples routinely to make sure that you're familiar with the participants' voices.
    a) If the interaction is not between the couple, you've completed the first pass for this audio file. Make the appropriate notation in the Audio Processing Logbook and move on to the next audio file.
    b) If the interaction is between the couple, you will need to ascertain whether the interaction is "interpersonal emotion regulation-relevant" (see below for more on the operational definition of IER-relevance). If the answer is "Yes", then add a new row to the **Interaction Record**: Note the filename, the start time of the interaction, and the end time of the interaction.
        a. **Note:** determining the start and end time may require listening to the prior audio recording, the subsequent audio recording, and/or listening to relevant section of the paired audio recording (e.g., Partner B's recordings). You should always listen for at least two minutes *after* you think that an interaction has ended to ensure that it doesn't continue after a brief hiatus or interruption.
        b. **On the Operational Definition of "IER-Relevant" Interactions**: IER-relevant interactions are moments when one or both partners are experiencing or have experienced an emotion that creates an opportunity for the other partner to manage or regulate those emotions. Occasionally, it may be very apparent that an interaction is IER-relevant. For example, one partner might say something like, "I had a really frustrating day at work today; do you mind if I tell you about it?" or "You're looking really frustrated; is there something I can do to help?" More often, you'll need to rely partially or entirely on more subtle cues. For example, instead of using an emotion word (like "frustrated"), a participant might instead say something like, "I feel shitty" or "I'm so stressed out right now." Likewise, a participant might ask for a hug because they've had a hard day or they might say that they "just need to vent." Or, they may not say anything about their affect/emotions at all, but you infer from the fact that they are yelling (likely anger) or sobbing (likely sadness) or simply the tone of their voice that they're experiencing an emotion. Remember, we're only interested in experiences of emotion that occur *when their partner is present* and so has the opportunity to respond. Be particularly on the lookout for any times that either partner describes an emotional experience, uses an emotion word to describe themselves or their

partner (including emotion-relevant behaviors, like "crying"), asks for or offers emotional support (e.g., comfort, validation, distraction from how they're feeling, etc.), as these are all indicators that an interaction may be IER-relevant.

i. **Note:** Not every IER-relevant moment will lead to what we might think of as a "whole interaction" with a lot of back-and-forth. For example, one partner may solicit or offer IER, but that solicitation or offer may be rejected or go unnoticed, in which case the "whole interaction" may only consist of a sentence or two!

ii. **Note:** IER-relevant interactions may not always occur face-to-face. For example, you might overhear a conversation between the couple that is IER-relevant.

c. If you have time remaining in your shift, you can go ahead and do the second pass (see below). This will involve listening to the same interaction from start to finish and making various ratings.

Second Pass(es):

a) Refer to the **Interaction Record**. Find the first interaction that has not already been double coded (i.e., undergone the second and third passes). Find the audio recording(s) that corresponds to that interaction and navigate to the interaction's start time as listed in the Interaction Record.
b) Listen to the interaction from start to finish. You should complete *at least* two full listens, potentially supplemented by additional reviews of dense or tricky portions. Keep detailed notes as you listen.
c) Open a link to the **Coding Workbook** (a Qualtrics survey) and code all relevant variables (see below).
d) **Before we describe the individual codes, there are a few general, but important notes about making content ratings:**
    i. Be *very* careful not to confuse your own reactions to situations and interactions with your evaluation of how the participants are reacting to those situations and interactions. For example, a participant might tell a story about a time that they made a mistake while giving a public speech. If you're particularly attuned to social-evaluative threat, you might be inclined to give the receiver a very high emotional intensity rating because you yourself would find the same situation highly anxiety inducing; however, the participant might have found the situation only mildly embarrassing. Or, a provider might say something that you personally find offensive, so you might be inclined to rate them as harsh and insensitive; however, the receiver might respond very positively (perhaps it's an inside joke).
    ii. Similarly, unless otherwise specified, codes should reflect *the interaction as a whole*, not just a portion of the interaction. If, during one brief portion of a much longer interaction, a provider says something that is harsh (possibly even very harsh), but it's really the only time in the whole interaction that they are at all harsh, your rating should reflect the presence of this harshness, but also the fact

that there was only a single harsh utterance across the interactions. In other words, don't confuse single moments with the overall frequency of a behavior.

i. **Caveat: A word about severity vs. frequency.** Many of our codes ask you to judge whether, e.g., one or both partners did "a little" of a behavior, "some" of a behavior, or "a lot" of a behavior. This is not strictly about how many utterances match that behavior, <u>although this is the single most important factor</u>. In making your rating, you should also think about the "severity" of the behavior. Our rule of thumb is that you can alter your final rating up or down by one category based on the severity of the utterances (e.g., to distinguish a series of mildly critical remarks from a series of highly derogatory remarks).

iii. Remember to pay attention to all of the various sources of information that you have at your disposal. Audio recordings are rich with information. It's not just about the words that our participants say, but how they say them (e.g., tone of voice), the care that they take in timing what they say (e.g., are they interrupting their partner, or are they waiting their turn?), how their partners react, and all of the other overt and subtle non-verbal cues (e.g., laughing, crying, sighing).

i. **Caveat:** In some cases, these various "information channels" may seem to disagree with one another. For example, someone could say that they're "screaming mad" in an apparently modest tone of voice. It's <u>not</u> your job to determine the ground level "truth"; it <u>is</u> your job to make the best summary rating possible based on your judgment and all of the available information.

iv. Ratings should always be independent of one another (see Individual Content Codes below). For example, you might think that if a provider was harsh, they can't possibly have also been empathic. This isn't necessarily the case, though; a provider can be harsh at one point during an interaction and empathic at a subsequent point. Therefore, to the greatest extent possible, you should be making a rating about harshness and then making an entirely separate rating of empathy. At the same time, a single utterance can encode multiple behaviors.

v. Always try to make a rating for each code. However, if you feel like you cannot make a particular rating, don't. Instead, indicate that you were unable to rate. **Note:** Rating these codes may require listening to a recording (or portions of a recording) multiple times to make sure that you've really understood the interaction or it might require listening to both Partner A and Partner B's tapes in case it's not clear from Partner A's recordings what Partner B is saying. Listen to the tapes enough times to be confident in your ratings (or confident that you cannot reasonably make a rating).

vi. Use your best judgment :)

Individual Content Codes

1. IER [0 = No; 1 = Yes]
   a) All of the interactions that appear in the **Interaction Record** have already been designated as "IER-relevant" (see operational definition above under "First Pass"). ***However***, <u>you may not agree that the interaction that you just listened to is IER-relevant</u>! **<u>This is OK.</u>** We are only interested in interactions that we can all agree really are

indicative of how our participants regulate one another's emotions. If you think that the interaction you just listened to is not IER-relevant for whatever reason, then mark "0" and do not code any of the other variables.

2. IER Solicited? [1 = Solicited, Explicit; 2 = Solicited, Implicit; 3 = IER not Solicited; 9 = Unclear]
   a) In some cases, the receiver will have solicited IER. This can take one of two forms: "explicit" solicitation or "implicit" solicitation. Explicit solicitation is the clearest case: it involves the solicitor explicitly asking their partner for emotion regulation or emotional support of one form or another. For example, the solicitor might say, "I had a really frustrating day at work; can I vent for a few minutes?" or simply "Can I vent for a few minutes?" Implicit solicitation is any time that the solicitor brings a negative emotion into the room, but doesn't explicitly ask for help. For our purposes, for example, saying "Ugh, I just had the most frustrating day at work!" is considered an implied opportunity for the provider to offer emotion regulation. In the case of implicit solicitation, it does <u>not</u> need to be clear that the emotion that the receiver is expressing is directed at the provider or that the would-be receiver is actually wanting regulation or support. For example, the receiver might slam a door: regardless of whether the receiver "intended" their partner to hear them slam the door, this could constitute an implicit IER solicitation. For more examples of what IER might "sound like," refer to the operational definition of IER (above).
   b) Sometimes, you may encounter a third scenario: a participant might offer IER to their partner without any apparent solicitation on the partner's part. For example, a participant might say something along the lines of, "You know, I've noticed that you've been really stressed about work recently; do you want to talk about it?" apropos of absolutely nothing that their partner has done or said (as far as you can tell from the recording). In these cases, mark 3 for "IER not solicited."
3. IER Offered? [1 = IER offered; 2 = IER not offered; 9 = N/A]
   a) Sometimes, one participant might solicit IER, but their partner might not respond with an offer of IER. This could occur when IER is solicited "implicitly," but there may also be cases where a more explicit ask goes unnoticed, ignored, or rejected by the person being asked for regulation. For example, the partner being solicited might have been distracted or might have simply not understood what the partner doing the soliciting was asking for.
   b) **What if they don't say "no," but they don't say "yes" either?** In some cases, it may be clear that a participant isn't offering IER even though they're not explicitly saying "no." For example, that partner might immediately change the topic or simply leave the room altogether. In other cases, though, it may be less apparent. For example, following an explicit solicitation by Partner A, Partner B may simply remain silent or they might say "mhm" in a manner that strikes you as so perfunctory that it really shouldn't count as a "yes"—for example, perhaps the participants had previously been talking about their plans for the day, and Partner B had spent *that* whole conversation simply mhm-ing. Be sure to listen to the whole interaction, however long it lasts. If at some point during the interaction, Partner B adopts a less passive stance and becomes actively engaged in the interaction (e.g., offers a hug, makes caring statements), that would constitute a 1 (IER offered). In general, however, err on the side of marking 2 (IER not offered) in the total absence of any engagement on Partner B's part.

c) **What if IER is offered, but not until later?** You could encounter a case in which Partner A solicits IER and Partner B responds by saying something like, "This seems like a really important conversation to have, but I'm actually in the middle of work; can we talk later please?" In these cases, mark 2 (IER not offered) *unless* Partner B subsequently reverses course and decides to offer IER *in the course of the same interaction*, in which case mark 1 (IER offered). Hopefully, if an IER interaction does subsequently take place, we'll catch it when it does and be able to code it separately.

d) **A word on Harsh Responding:** This variable should be coded *independently* of Harsh Responding (see below). It is true that, in some cases, a participant may reject an IER solicitation in a manner that is <u>actively</u> critical or dismissive (e.g., "Honestly, honey, I just can't be bothered with your problems right now"), in which case you would mark 2 (IER not offered) and also make an appropriately high rating for Harsh Responding. In other cases, however, as in many cases in which an implicit solicitation may simply go undetected by a participant, you might mark 2 (IER not offered), but you wouldn't indicate that the provider was actively critical or dismissive.

e) **Note:** If Code 2 is a 3 (i.e., IER not solicited) this should (almost) always be a 1. That is, if you've determined that this is an IER-relevant interaction and it wasn't solicited by the receiver, then it must have been offered without solicitation.

f) **Note:** In cases in which IER is solicited, but not offered, you may not be able to make all (or even most) of the content ratings below. **As always, only make ratings that you are confident about.**

4. IER Offer Accepted/Rejected [1 = IER accepted; 2 = IER rejected, ignored, not noticed, or otherwise not accepted; 9 = Unclear]

a) Just because Partner A offers interpersonal regulation, doesn't mean that Partner B either notices or accepts the offer. For example, Partner B might respond to an offer of interpersonal regulation by saying, "I'm fine" or "I don't want to talk about it." These are examples of an IER offer being actively rejected; if that's the end of the interaction, code as a 2. However, if IER is *initially* rejected, but is subsequently accepted in the course of a single interaction, code that interaction as a 1. We would also code as a 2 situations in which IER offers aren't actively *rejected*, but are ignored or seem to go unnoticed (e.g., Partner B changes the topic, storms out of the room without saying another word, etc.).

b) **Note:** This code should be rated even when the IER is solicited, especially when it is solicited implicitly. For example, we can imagine a case in which one partner comes home and immediately slams the door in anger, but doesn't ask for help (Implicit Solicitation; Code 2 rated as a 2), which is followed by an offer (Code 3 rated as a 1), that is summarily rejected (Code 4 rated as a 2).

c) **Note:** In cases in which IER is offered, but ignored or rejected, you may not be able to make all (or even most) of the content ratings below. **As always, only make ratings that you are confident about.**

5. Is Partner A or Partner B the provider of interpersonal regulation? [1 = Partner A is the Provider; 2 = Partner B is the Provider; 3 = The interaction is mutually regulatory]

a) In some cases, it may be very clear that one partner is the provider of IER and the other partner is the receiver. For example, one of the participants may have had a frustrating interaction with a friend, and their partner may offer them support. In other cases, both participants may be experiencing an emotion that creates opportunities for regulation on both sides. For example, they might both be worried about their children, and be offering

one another support. If it's the case that the interaction is mutually regulatory, you will need to make receiver **and** provider content ratings for both participants (see below).

a. If the interaction is mutually regulatory, consult the **Interaction Record**. Add a new row with a new Interaction ID to reflect the "other half" of the interaction if one does not already exist (see below).
    i. In the appropriate column in the Interaction Record, note that both the new interaction you've just added to the Record and the interaction that is it's "other half" are mutual regulation interactions. Example: you're doing the first pass through Interaction 313 and you realize that it's a mutual regulation interaction. Open the **Interaction Record** and add a new row for Interaction 314. For interaction 313, input "Yes: 314" in the "Mutual Regulation?" column. For interaction 314, input "Yes: 313" in the "Mutual Regulation?" column.
    ii. Additionally, under the header "Provider's ID" column, note whether Partner A (e.g., Participant 101A) or Partner B (e.g., Participant 101B) was the provider for this row. For example, if you're doing the first-pass through Interaction 313 and you've decided to focus on Partner A to answer the provider-specific codes (and Partner B to answer the receiver-specific codes), then you should list Partner A as the provider for Record 313 *and* you should indicate that Partner B is the provider for Record 314. **Be VERY careful that you're coding the correct partner when you're coding a mutually regulatory interaction. If you're the first person to code either "half" of the interaction, then you can pick which partner you want to code. Otherwise, the Interaction Record should indicate whether Partner A or Partner B is the provider for that record.**
    iii. Note: If you're doing the first-pass through Interaction 314, you won't have to add another row to the Interaction Record, since Interaction 313 is already in the Record. This will be obvious, since both 313 and 314 will already be marked as yoked interactions.

b. You'll need to fill out the Coding Workbook survey twice: once for each "half" of the interaction (i.e., once for Interaction 313 and once for interaction 314).

6. Is the locus of the interaction primarily internal or primarily external to the relationship? [1 = Internal, and the cause of the emotions was the provider; 2 = Internal, but the cause of the emotions was not the provider; 3 = External; 9 = Unable to make a rating]

a) Sometimes, our partners do things that make us frustrated or upset. Or, we may be worried about the future of our relationship for whatever reason. If we then engage in an interpersonal emotion regulation interaction with our partner, the locus of the interaction is *internal* to our relationship, meaning that the person *providing* regulation is implicated in the cause of our emotions. This would also include scenarios that are about the family as a whole (e.g., conversations about finances, parenting, etc.). When the locus of the interaction is internal to the relationship, we further distinguish between cases in which the partner is the direct cause of the emotion (e.g., "I'm just so mad at you because…"; coded as a 1) and cases in which the partner is implicated, but not the cause of the emotion (e.g., "I got this letter from one of Sara's teachers, and now I'm really concerned about our daughter"; coded as a 2). However, there are also plenty of other times when the cause of our emotions has little or nothing to do with our partner or our relationship.

For example, we might have had a frustrating day at work, and want to share that experience with our partner. In this case, the locus of the interaction is primarily *external* to the relationship (coded as a 3).

b) You might occasionally encounter a corner case in which the locus of the interaction seems to be simultaneously internal and external. For example, a participant might relate to their partner a frustrating experience at work (primarily external) that has them worried about the family's finances (primarily internal). Use your best judgment, but err on the side of determining that the locus is internal if the provider is implicated in any way in the emotional experience (e.g., in this case, the interaction might trigger the provider's own worries about the family's finances, which implicates them in the interaction).

7. Receiver Primary Emotion [1 = Sad/Unhappy; 2 = Anxious/Worried/Afraid; 3 = Angry/Frustrated; 4 = Self-Conscious Emotions (Shame/Guilt/Embarrassment); 5 = Unspecified Negative Emotion (e.g., "upset", "distressed", "stressed", "bummed"); 6 = Happiness/Any Positive Emotion; 9 = Unclear/Too Mixed to Distinguish/Other Specified Emotion]
   a) We are interested in which emotions receivers are experiencing immediately preceding and during IER interactions that they are seeking to manage or regulate. We refer to these emotions as the "primary emotions" or the "target emotions." You may be able to infer the target emotion from direct and indirect statements by one or both partners or by the IER receiver's behavior. In some cases, partners may refer to various specific emotions throughout an interaction, but you can nevertheless discern that one specific emotion is clearly the primary or target emotion. In other cases, however, participants may have considerable difficulty labeling their emotions or may be experiencing many emotions simultaneously (referred to as a "mixed" emotional state); if you're not able to make a confident inference, fall back on 5 (Unspecified Negative Emotion) or 9 (Unclear/Too Mixed to Distinguish/Other Specified Emotion).
   b) **Direct statements:** IER receivers may sometimes use emotion words to describe how they're feeling. For example, they may say that they're feeling irritated or sad.
   c) **Indirect statements, contextual clues, and tone of voice:** On other occasions, IER receivers may never use a specific emotion word to describe how they're feeling. For example, they might say, "I'm just feeling really shitty." In these cases, we may still be able to infer the target emotion from indirect statements, contextual clues, and tone of voice. For example, the receiver might say, "I feel like punching something" (likely anger) or the provider might say, "You're crying."
   d) **A word on synonyms:** Occasionally, someone might say something like, "I'm not angry; I'm just annoyed!" On the one hand, they're denying being angry; on the other hand, they're using a synonym for anger. For our purposes, we would (most likely) describe the primary emotion as "Angry/Frustrated" in an instance like this one. There are many other synonyms for anger and frustration (such as mad, irked, vexed, annoyed, exasperated, furious, outraged, enraged, irate, irritated), sadness (such as depressed, dejected, blue, melancholy), anxious/worried/afraid (such as concerned, nervous, scared, uneasy, frightened), and ashamed/guilty/embarrassed (such as apologetic, contrite, humiliated, regretful, remorseful). Be on the lookout for synonyms like these, not just the primary emotion labels that we are using.
   e) **Distinguishing between specific emotions:** Although we rely substantially on lexical and colloquial definitions of specific emotions, we also include a brief note on emotion

theory to aid in distinguishing between specific emotional experiences. To be clear, however, an emotion that you encounter on one of the recordings does not need to exactly fit the prototypical definition of an emotion to be labeled as such.

a. **Sadness.** Sadness is typified by a negatively valenced response to loss or grief and is characterized by feelings of hopelessness, helplessness, and dejection. Sadness is often accompanied by crying, withdrawal, and lethargy.
b. **Anxiety.** Anxiety is typified by concern and negative feeling about future events. It is, in other words, a *future-oriented* mood state in which one perceives oneself as not being prepared to attempt to cope with upcoming negative events. It is frequently accompanied by worry and catastrophizing, cognitive patterns that emphasize negative anticipation of the feared event. For our purposes, we are lumping together anxiety and fear, the latter of which is the affective response to a currently present or recently present threat (e.g., "I just saw a snake and it freaked me out!"). Anxiety is often accompanied by restlessness, feelings of stress and dread, and (muscle) tension. Anxiety is likely to prompt individuals to want to avoid the upcoming event that is causing them to be anxious.
c. **Anger.** Anger and frustration are negatively valenced responses to perceived provocations, slights, hurts, setbacks, or threats that prompt us to want to fight back against an individual, group, or system. We may become angry (at others or at ourselves) when we perceive ourselves as having been thwarted in our efforts to achieve a goal or as making insufficient progress toward a goal. Unlike sadness and anxiety, anger typically prompts us to want to approach whomever it is that is making us angry in order to retaliate or otherwise rectify the situation. Anger is frequently accompanied by yelling, (muscle) tension, and physiological arousal.
d. **Self-Conscious Emotions.** To some extent, all emotions (including sadness, anxiety, and anger) involve an element of self-consciousness; for our purposes, however, we refer to the specific emotions of shame, guilt, and embarrassment as primarily "self-conscious emotions." These are negative emotions that threaten our self-concept and that cause us to be concerned about how other members of our community will perceive us. These emotions are associated with the subjective experience that we are or will be perceived by others to be worthless, broken, failures, or otherwise inferior (e.g., "I'm a terrible person." "There's something terribly wrong with me."); the sense that we've done something wrong or that we have violated our own values or the values of our society (often accompanied by feelings of remorse or regent); and the sense that we have made ourselves vulnerable to negative social evaluation or humiliation. These emotions may variously prompt us to withdraw or hide ourselves from social situations, to apologize for ourselves or our behavior, or to engage in self-deprecation.

8. Receiver Emotional Intensity at the Start of the Interaction [1-7 Likert scale, where 1 = Not very intense at all and 7 = Extremely intense; 9 = unable to make a rating]
    a) Direct statements, but usually indirect clues: Occasionally, an IER receiver might give a direct indication of the intensity of their emotions. For example, they might say something like, "This is the saddest I've felt in forever." More commonly, however, we can infer something about the intensity of their emotions from indirect clues. For example, they or their partner might reflect on a behavioral clue (e.g., crying,

trembling). There might be clues in the way that they're speaking; for example, speaking unusually rapidly or loudly when feeling angry.

9. Receiver Intensity of the Target Emotion at the End of the Interaction [1-7 Likert scale, where 1 = Not very intense at all and 7 = Extremely intense; 9 = unable to make a rating]
    a) See above.
10. Receiver Overall Interaction Satisfaction [1-7 Likert scale, where 1 = Very dissatisfied, 4 = Neither satisfied nor dissatisfied, and 7 = Very satisfied; 9 = unable to make a rating]
    a) Whether the receiver is satisfied with the interaction or not is not necessarily the same thing as whether their emotional intensity decreased over the course of the interaction. For example, a participant might wear themselves out screaming invective at their partner (and so the intensity of their anger might diminish), but, at the end of that same interaction, they might say something like, "You know, I was really hoping for more from you" or "It just seems like we're never on the same page when we get angry." Statements like these would indicate a low level of satisfaction, independent of the change in emotional intensity. In essence, you're looking to make a rating of the overall degree to which the receiver walked away from the interaction feeling as though it "went well" or "was a good interaction." This means that you should focus primarily on where the receiver is at the end of the interaction. For example, they might have seemed relatively dissatisfied at the midpoint of the interaction; however, if by the end of the interaction they appear to be relatively satisfied, your rating should reflect that the interaction was ultimately a satisfying one.
11. Provider Overall Interaction Satisfaction [1-7 Likert scale, where 1 = Very dissatisfied, 4 = Neither satisfied nor dissatisfied, and 7 = Very satisfied; 9 = unable to make a rating]
    a) Just because the receiver is satisfied and/or the provider is responsive does not necessarily mean that that the provider is satisfied with the interaction and vice versa. As with receiver satisfaction, you're trying to make an inference about whether the provider felt as though the interaction "went well."
12. Provider Overall Attunement [1-5 Likert scale, where 1 = 0% Attuned, 3 = 50% Attuned, and 5 = 100% Attuned; or code 9 if you were unable to adequately gauge the provider's attuned, e.g., because the receiver wasn't sufficiently clear about what it was that the receiver wanted or didn't want]
    a) There's frequently a difference between what our partners want us or ask us to do to support them and what we actually do. This is what we mean by "attunement." If we ask our partners just to let us vent to them and not to try to solve the problem, but they repeatedly interrupt us with suggestions, we would describe this as low attunement.
    b) A single interaction may include multiple asks and multiple responses. Be sure to take this into account. For example, a provider may be *un*attuned to the receiver's initial request to let the receiver speak without interrupting, but, later on, they may be attuned to the receiver's request to help with problem-solving. Remember that you are rating the overall rate of attunement, not their degree of attunement to any single request.
    c) Crucially, attunement is distinct from satisfaction (see above) or effectiveness. Our partners can do *exactly* what we ask of them, and we can still wind up feeling dissatisfied by the interaction. Alternatively, we can ask them not to offer problem-solving, but be pleasantly surprised when the problem-solving that they offer turns out to be genuinely helpful (low attunement, high satisfaction).

d) Frequently, you will simply not be able to rate attunement because the receiver will not be explicit about what they want (or don't want) from the provider. We are intentionally setting a high bar here; avoid attempting to infer what the receiver wants.

13. Provider Empathic Responding [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) For our purposes, empathic responding is defined in terms of understanding and validating the partner's thoughts and feelings. This encompasses not only *attentiveness*, but also *responsiveness*. Thus, empathic responding is about being there **with** someone *and* being there **for** them: they're in this together and they're giving the receiver a space in which to feel heard and understood. They feel what their partners feel and their partner's emotion have a strong effect on their own emotions. For example, the provider may find more or less subtle ways to let the receiver know that their feelings are understandable and legitimate; the provider may label the receiver's emotions so that they receiver feels understood; the provider may ask directly about the receiver's feelings or otherwise convey that they're taking a sincere interest in the receiver's feelings, or the provider may indicate that they share in or reciprocate the receiver's feelings. In addition to direct statements of understanding (e.g., "Yeah, that makes a lot of sense"), this might also include more subtle indicators (e.g., "Mhm, mhm"). However, be wary of perfunctory behaviors: if the provider never contributes more to the interaction than the occasional distracted "mhm," they may in fact be mostly or entirely disengaged. For example, if all a provider is contributing is the occasional "mhm," you might hear a receiver check in to see if their partner is really listening. Evidence that the provider is distracted or disengaged or that the receiver is skeptical that the provider is really paying attention or understanding what is being said should contribute to a correspondingly low rating for Empathic Responding.
14. Provider Harsh Responding [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) Our partners may not always respond positively in the context of IER. For example, a provider might be generally uncomfortable with emotionality or think that the emotion is unwarranted in a particular situation. They may have limited emotional bandwidth themselves. Whatever the reason, harsh responses can occur during interpersonal regulation interactions and may take various forms, including criticizing, blaming, and insulting the receiver; dismissing and ignoring the receiver; invalidating the receiver; expressing anger, frustration, or contempt at the receiver; and engaging in negative humor, mockery, or sarcasm (but not, e.g., playful, affiliative teasing).
    b) **What about "tough love"?**: Sometimes, a participant might think that they need to say harsh things to their partner *because* they love them. We sometimes refer to behaviors like this as "tough love." And, we may not always be able to distinguish between tough love and plain meanness. That's OK. If a harsh statement is tempered by a loving one (e.g., "You know I'm only saying this because I want to see you succeed"), you can take this into account in making your rating (see note above re: "severity"), but harsh statements can still be harsh, even when they take the form of "tough love" and so should be coded as such. Recall that you will also be providing a separate rating for Warmth (see below).

15. Provider Offered, Provided, or Encouraged Positive Distraction [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) Perhaps the provider suggested that the receiver take a walk, listen to music, or engage in some other pleasurable activity to take their mind off of their emotion. Or maybe they tried to change the topic to a more pleasant one (but not in a domineering, defensive, or dismissive way). These are positive examples of offering or providing a distraction. The goal here is not to resolve the emotion-inducing situation, but to get some mental, physical, or temporal distance from it.
16. Provider Expressed Confidence in the Receiver or in the Strength of the Relationship [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) Sometimes, what we want to hear from our partner in an emotional situation is that they're confident that we're going to be OK because we've proven our resilience, our intelligence, or our savviness in the past, or simply that they have confidence in us. **Importantly, these statements contain an element of positive regard and praise that goes beyond simple reassurance**—that is, expressions of confidence go beyond saying "It's all going to be OK." As such, be careful to distinguish utterances that contain expressions of confidence from the next code, Reassurance.
    b) Note that this does <u>not</u> include expressions of confidence that are unrelated to the receiver or the relationship; for example, a statement like, "I'm sure that everything is going to be all right because your friend Jim is a forgiving person" would not be incorporated in this rating, but would instead be incorporated in Reassurance.
17. Provider Offers Reassurance [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) Any vague or specific reassurances that "Everything is going to be OK" or similar **that aren't counted toward expressions of confidence (see above)** should be counted toward general reassurance. <u>Avoid double-counting instances of Expressions of Confidence and instances of Reassurance.</u>
18. Provider Practical Responding [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) Providers frequently respond to emotional situations and/or requests for emotional support by focusing on pragmatics and on changing or "fixing" the emotion-inducing situation. This could take the form of offering advice, prompting the receiver to engage in problem-solving, providing relevant factual information, or similar.
    b) One way to think about Practical Responding is to think about the nature of the external situation that is prompting the receiver's emotions. For example, the external situation might be an impending work deadline. If the provider offers advice or problem-solving that is focused on altering or responding to that external reality (e.g., "Let's make a plan for meeting that deadline"), then they're problem engaged in some form of Practical Responding.
    c) Be particularly careful to distinguish Practical Responding from the next code, Cognitive Responding.

19. Provider Cognitive Responding [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) There's a subtle, but important distinction between cognitive responding and practical responding. Practical responding is all changing the external reality by problem-solving. In contrast, cognitive responding is all about changing how the receiver is *perceiving* their situation (opposed to actually changing the situation itself). This might look like suggesting alternative interpretations of the situation to help their partner see the situation or their feelings in a new light (e.g., "maybe that other driver was rushing to their spouse to the hospital, and that's why they cut you off"). In essence, the provider is prompting or encouraging the receiver to adopt a different perspective, potentially by offering the provider's own perspective. It could also be by asking questions that prompt the receiver to reconsider the way they are perceiving a situation or by responding positively to the receiver's own efforts to reappraise their situation.
    b) Take the example of the impending work deadline (see Code 19). Instead of helping the receiver to make a plan, the provider might suggest that the receiver's project is a really great opportunity to showcase their hard work. This suggestion does not in any way alter the external situation, but it might affect how the receiver is viewing the situation (i.e., as a challenge / opportunity for growth rather than as a threat).
20. Provider Warmth and Affection [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) In essence, you're looking for times when the provider does or says something that is clearly intended to make the receiver crack a smile or laugh, that makes the receiver experience the provider's warm regard for the receiver, or that is intended to express the depth of caring or love.
    b) Specific examples of warmth and affection behaviors might include:
        a. Positive Humor
        b. Caring Statements, such "I love you," "I care about you," "I worry about you," and so forth.
        c. Positive Reminiscing
21. Provider Overwhelmed [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) Sometimes, a provider may be overwhelmed by the receiver's emotions. The feeling of being overwhelmed *might* cause the provider to react harshly (in which case their harshness should be reflected in your rating of Code 14), *but* being overwhelmed and reacting harshly are distinct entities. For example, you can imagine a provider saying statements like, "This is all a bit much for me to process" or (somewhat more obliquely) "I don't honestly know what to say" in a tone of voice that conveys that they're genuinely trying to understand (high Empathic engagement, low Harsh Responding), but are simply out of their depth (high Overwhelmed).
22. Provider Defensiveness & Other Egocentric Behaviors [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]

a) Defensiveness functions to deflect responsibility or blame. It communicates a kind of innocent victimhood or righteous indignation (e.g., as a counterattack) on the part of the speaker, implying that whatever bad thing being discussed is not the speaker's fault. Crucially, defensiveness (and other related behaviors) turn the interaction from one that's primarily focused on the needs of the receiver into one that's primarily focused on the needs of the provider: you thought this was about you, but actually it's about me. For example, in an IER-relevant interaction that is internal to the relationship and where the cause of the would-be receiver's emotions is the provider, the provider might perceive themselves as being criticized or attacked and thus be motivated to defend themselves.

23. Provider Expressive Suppression [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) For our purposes, expressive suppression refers to any statements made by the provider that indicates that the receiver needs to (or ought to or should) "contain" their emotions or displays of emotions. This is different than invalidation (in which the provider communicates that the receiver "shouldn't be feeling the emotions that they're feeling") and needn't (necessarily) be harsh. For example, a provider might say, "Quit crying!; we're in public, and people are looking at us" (harsh and suppressing) or they might say, "You just need to hold it together for a little longer and then we'll have gotten past this" (suppressing, but less harsh). Or they might say, "You're making me nervous; you need to chill out" (suppressing and also overwhelmed).
24. Provider Acceptance [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) For our purposes, acceptance is ***not*** merely the absence of an active attempt by the provider to change how the receiver is feeling. It's also not empathy or validation (e.g., statements that communicate that the receiver's emotions are understandable). Rather, acceptance refers to statements that actively communicate the provider's belief that the receiver does not need to change either themselves or their emotions. Likewise, Acceptance is distinct from Expressions of Confidence (Code 16). For example, the statement "I don't need you to be less sad; I just need you" would be high Acceptance and high Warmth (Code 20), but does not involve an explicit expression of confidence. On the other hand, the statement, "I don't need you to be less angry; in fact, it's your passion that makes me so confident that we're going to get through this together" would be high Acceptance, high Warmth, *and* high Expressions of Confidence.
    b) As another example, imagine that one of our participants has just learned that a family member has passed away. Their partner might tell them something like, "Grieving is a natural part of the process; we can be sad together." This utterance communicates that the partner's emotions are understandable (Empathy) *and* do not need to change (Acceptance).
    c) Another situation when Acceptance might frequently be relevant is anytime when the receiver apologizes for the emotions that they are experiencing or expressing. If the provider responds by communicating that there is no need for an apology, that would constitute a classic acceptance utterance (and potentially also Empathy).

25. Co-Brooding [1 = No, they didn't do any of this; 2 = Yes, they did a little of this; 3 = Yes, they did some of this; 4 = Yes, they did a lot of this; 5 = Yes, this was one of the main things they did; 9 = unable to make a rating]
    a) Sometimes, individuals or groups of individuals can become ruminatively focused on the details of a past event that caused them distress, going over and over every detail of the situation with their partner even after it's clear that the partner understands the pertinent details (e.g., after telling the provider that a friend yelled at them, the receiver might say three or four more times, "I just can't believe they said that to me!" over the course of the conversation). Co-brooding is about focusing on the negatives. Receivers and providers alike may engage in bitter or despondent recrimination ("Why couldn't I/you have just handled this situation better?" "Why do I/you always react this way?" "I just wish it had gone better" "What did I do to deserve this?"). These patterns aren't really about generating novel solutions or genuine efforts to reflect on the situation (contrast Cognitive Responding); in fact, they're less about looking forward to the future than they are about being stuck in the past. When couples are co-brooding, they're stuck in a loop, reiterating their grievances (whether with themselves or others) over and over again.
    b) **Note:** This code is unique in that it asks you to make a rating that takes into account what both partners are doing simultaneously. Rather than trying to weigh each of their separate contributions, you should instead be thinking about the fraction of the interaction in which one or both partners seem to be "stuck" in this brooding loop. For example, if the entire interaction is dominated by co-brooding, you should rate this a 3, even if the provider isn't doing anything apparently to encourage the brooding.
26. Intimate Partner Violence [1 = No, there was no evidence of intimate partner violence; 2 = Yes, there was evidence of intimate partner violence]
    a) Separate from IER, we are also interested in experiences of intimate partner violence (intimate partner violence). As part of our initial pass through these audio recordings, we will not be coding all of the various components of intimate partner violence; however, we do want to make a note if any of the recordings that we are coding seem relevant to intimate partner violence. For our purposes, code 2 (evidence of intimate partner violence) if you overhear any instances of participants bullying, threatening, insulting, humiliating, belittling, screaming or cursing at, or physically injuring their partners.