

# Relative Importance of Conflict Topics for Within-Couple Tests: The Case of Demand/Withdraw Interaction

Alexander O. Crenshaw and Karena Leo  
University of Utah

Andrew Christensen  
University of California, Los Angeles

Jasara N. Hogan, Katherine J. W. Baucom, and Brian R. W. Baucom  
University of Utah

Researchers commonly employ observational methods, in which partners discuss topics of concern to them, to test gender differences and other within-couple differences in couple conflict behavior. We describe a previously unidentified assumption upon which statistical tests in these observational studies are frequently reliant: whether each partner is more concerned or dissatisfied with the topic selected for them than the partner is. We term this the *relative importance assumption* and show that common procedures for selecting conflict discussion topics can lead to widespread violations of the assumption in empirical studies. Study 1 conducts a systematic review of the literature and finds that few existing studies ensure relative importance is met. Study 2 uses two empirical samples to estimate how often relative importance is violated when not ensured, finding it is violated in one third of interaction tasks. Study 3 examines the potential consequences of violating the relative importance assumption when testing within-couple differences in observed behavior, focusing on gender differences in the demand/withdraw pattern. Results show that these tests were profoundly impacted by violations of relative importance. In light of these violations, we conduct a more rigorous test of demand/withdraw theories and clarify previously inconsistent results in the literature. We recommend explicit consideration of relative importance for studies testing within-couple effects, provide methodological recommendations for selecting topics in future studies, and discuss implications for clinical practice.




**Keywords:** intimate relationships, communication, conflict, observational coding, demand/withdraw

**Supplemental materials:** <http://dx.doi.org/10.1037/fam0000782.supp>

Observational methods are commonly used to understand gender differences and other within-couple differences in couple conflict behavior. These methods have been essential tools for understanding behavior in intimate relationships, especially for identifying and explaining gender differences (e.g., Heavey, Layne, & Christensen, 1993). The selection of conversation topics has previously been identified as an important, but at times underappreciated, decision that can impact conclusions drawn from observational studies, such as the positivity or negativity of an interaction or the impact of gender (Heyman, 2001). The present

investigation builds on this work by describing a previously unidentified assumption upon which statistical tests of within-couple variables are often dependent: that each partner is more dissatisfied about their chosen topic than the other partner is about that same topic. In a systematic literature review and two empirical studies with 182 couples, we a) explore how frequently the literature presently attends to this issue, b) estimate how often the assumption may be violated when not attended to, c) show how assumption violations can impact the validity of study results for tests of within-couple variables, and d) in light of assumption violations,

This article was published Online First July 30, 2020.

 Alexander O. Crenshaw and  Karena Leo, Department of Psychology, University of Utah; Andrew Christensen, Department of Psychology, University of California, Los Angeles;  Jasara N. Hogan, Katherine J. W. Baucom, and Brian R. W. Baucom, Department of Psychology, University of Utah.

This article and data collection were supported in part by the Graduate Student Research Award awarded to Alexander O. Crenshaw by the Consortium for Families and Health Research (C-FAHR) at the University of Utah, discretionary funds provided to Brian R. W. Baucom by the University of Utah, grants from the National Institute of Mental Health awarded to Andrew Christensen at UCLA (MH56223) and Neil S. Jacobson at the University of Washington (MH56165), and a research career development award from the

National Institute of Diabetes and Digestive and Kidney Diseases awarded to Katherine J. W. Baucom (K23DK115820). We thank members of the CLOSE Lab at the University of Utah for feedback on earlier versions of this article. Study materials, including code for emulation procedures, can be found at <https://osf.io/gdp3e/>.

This article is based on data that were presented at the 2018 annual conference of the Association for Behavioral and Cognitive Therapies in Washington, DC.

Correspondence concerning this article should be addressed to Alexander O. Crenshaw, Department of Psychology, University of Utah, 380 South 1530 East BEH S 502, Salt Lake City, UT 84112. E-mail: [alexander.crenshaw@utah.edu](mailto:alexander.crenshaw@utah.edu)

revisit previously inconsistent results regarding three competing theories for an important marital interaction pattern. We close with methodological and reporting recommendations for future research.

### Topic Selection in Observational Studies of Couple Conflict Behavior

Selecting topics for partners to discuss is a key step in observational studies of intimate relationships. Studies about conflict may ask partners to discuss an unresolved relationship issue to elicit conflict behaviors (e.g., Gordon & Chen, 2016), whereas studies about social support might ask partners to discuss an aspect of themselves they want to change to elicit supportive behaviors (e.g., Girme, Overall, Simpson, & Fletcher, 2015). Along with deciding how to select the specific topics, researchers must also decide the number of discussion topics. Focusing on conflict, studies commonly ask partners to discuss either one, jointly decided topic (e.g., Lemay, 2014) or two topics, one for each partner (e.g., Heavey et al., 1993). Selecting two topics is advantageous by allowing each partner to discuss an issue of interest to them and is grounded in conceptual models suggesting behavior differs according to who raised the topic or how important the topic is to each partner (e.g., Heavey et al., 1993), and thus has been recommended over single-topic designs (Heyman, 2001). Using two topics in combination with added procedural control, such as randomizing and counterbalancing the order of conversation topics, also allows researchers to go beyond simple observation by unpacking the possible causes or drivers of behaviors that vary according to which spouse has selected the topic for discussion (e.g., Heavey et al., 1993).

However, when within-couple differences are of interest, statistical tests of their impact on behavior depend on a critical, previously overlooked assumption, which we term the *relative importance assumption* of the conflict interaction paradigm. Within-couple differences are those such as gender differences in mixed-gender couples (Heavey et al., 1993), culture in intercultural couples (Hiew, Halford, van de Vijver, & Liu, 2016), desire for change (Heyman, Hunt-Martorano, Malik, & Slep, 2009), income or employment differences (Cohen & Levin, 2012), or power (Sagrestano, Heavey, & Christensen, 1999). The *relative importance assumption* is that each partner is more dissatisfied with or wants more change about their topic than the other partner does about the topic. This assumption's importance is derived from the fact that whose topic is being discussed is a distinguishing factor, much like Kenny's description of distinguishable (vs. indistinguishable) dyads (Kenny, Kashy, & Cook, 2006). Distinguishable dyads are those in which members can be identified by a characteristic that differs within the dyad (e.g., gender in mixed-gender couples), whereas indistinguishable dyads have no such characteristic (e.g., gender in same-gender couples). However, conceptual distinguishability is not sufficient; distinguishability must also be established statistically (Kenny et al., 2006). When indistinguishable dyads are treated as distinguishable, estimates of associations can vary based on arbitrary group-assignment decisions (Woody & Sadler, 2005), biasing statistical results.

In two-topic designs, the distinguishing factor is which partner is more dissatisfied about the topic being discussed. If the relative importance assumption is not met in these designs, the topics

become indistinguishable because one partner may be more dissatisfied about both topics. In this case, effects of topic become experimentally confounded with other variables, such as gender. For example, a finding that women enact more demanding behaviors than men during conflict (e.g., Christensen & Heavey, 1990) could be due to gender differences in demanding (i.e., a true gender effect), or it could be because the female partners were more invested in both discussion topics that were selected (i.e., a topic effect disguised as a gender effect). If this were the case, our understanding about the causes of some marital conflict behaviors, such as gender differences in the demand/withdraw pattern (described in Study 3), may be inaccurate.

Unfortunately, common procedures, even those consistent with recommendations (Heyman, 2001), do not ensure that relative importance is met. Doing so would require comparing ratings *across* partners to ensure that partner 1's topic is rated higher by partner 1 than by partner 2, and vice versa for partner 2's topic. However, procedures typically allow partners to select their own topic, or they compare topics only *within* a person to choose each person's highest-rated topic.<sup>1</sup> As shown in Figure 1, this procedure allows for one partner to be more dissatisfied with both topics, violating relative importance. In this example, both partners are assigned their top-rated topics: "having interesting conversations" for the husband and "handling family finances" for the wife. However, the wife, who is more dissatisfied overall, ends up more dissatisfied with both topics to be discussed. A related problem emerges for single-conflict studies, which typically select topics by choosing the topic with the highest cumulative disagreement across both partners (e.g., Crenshaw, Leo, & Baucom, 2019) or by asking the couple to jointly decide on a topic (e.g., Lemay, 2014, Study 1). Neither procedure ensures the topic is of equal interest to both partners. If one partner is more dissatisfied with the issue, topic effects become confounded with other effects, like gender (Stanley, Bradbury, & Markman, 2000).

### Present Studies

The present investigation is comprised of three parts. Study 1 conducts a systematic review of the literature to identify how often observational studies of romantic partners attend to the issue of relative importance of topics. Study 2 estimates how often relative importance is violated when it is not ensured, and how often a related problem emerges in single-conflict studies. Study 3 tests if and to what extent relative importance assumption violations impact tests of within-couple differences in behavior, using gender differences in the demand/withdraw (D/W; Christensen & Heavey, 1990) interaction as an illustration. Study 3 also revisits three competing theories about the drivers of D/W behavior in light of

<sup>1</sup> An additional criterion used by some (B. Baucom et al., 2010; Heavey et al., 1993; Holtzworth-Munroe et al., 1998) is to require that selected topics are similarly rated (within 1 or 2 points) by their respective chooser. For example, if partner 1 rates their top topic a 7 (completely dissatisfied), they can only be assigned that topic if partner 2 also rates their top topic a 6 or 7. If that criterion is not met, the experimenter chooses a lower-rated topic until each person rates their topic similarly to the other. We reason that this procedure should reduce the number of relative importance assumption violations, but it does not ensure the assumption is met, and is unlikely to eliminate violations. Violations can still occur, for example, when one partner rates multiple topics the highest, and the other partner's top topic is also one of those topics.

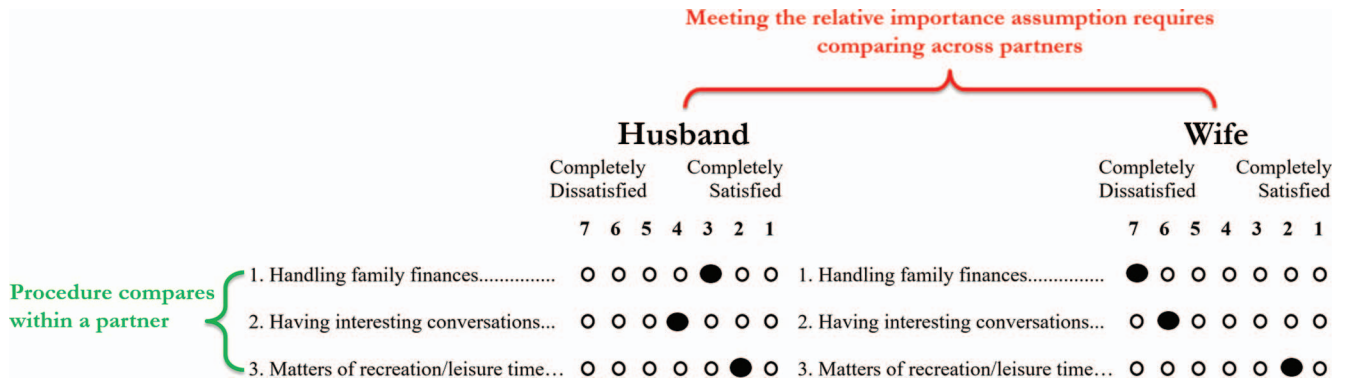


Figure 1. Example subset of a questionnaire used to select conversation topics (Heavey et al., 1995). The common method for selecting topics would assign “having interesting conversations” to the husband and “handling family finances” to the wife. However, in fact, the wife is more dissatisfied than the husband about both of these topics. See the online article for the color version of this figure.

assumption violations to clarify previously inconsistent results. Study materials can be found at <https://osf.io/gdp3el>.

### Study 1

The literature review focused on observational behavioral research of couple communication published between 2013 and 2019 in five major journals that commonly publish romantic relationship research: *Journal of Personality and Social Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Family Psychology*, *Personal Relationships*, and *Family Process*. We performed a keyword and abstract search in PsycINFO using the following terms: *observational coding*, *behavior*, *discussion*, *conflict*, *support*, *interaction*, *communication*, *couples*, *marital*, and *marriage*. Inclusion criteria were that the study coded observable behavior of one or more structured conflict discussions from video- or audio-recordings involving only romantic dyads<sup>2</sup> (i.e., excluding families and couple therapy sessions). For multistudy papers, we examined only the first study meeting criteria to avoid biasing results with multiple studies within a paper that are likely to use similar methods. We note when studies from different publications appeared to use the same sample and conversation(s) to avoid double counting.

The first author extracted the following information from each study: 1) number of structured conflict discussions; 2) for studies using the two-topic conflict paradigm, whether procedures ensured relative importance was met (i.e., compared both self and partner ratings of each topic to ensure the topic chooser was more dissatisfied or wanted more change); and, 3) for studies using a single conflict interaction, whether topic selection ensured the topic was equally important to both partners (i.e., compared partners’ ratings to ensure the selected topic was equally rated by both partners). We anticipated that studies would rarely ensure the assumption is met, but otherwise approached Study 1 in an exploratory manner.

In total, 49 studies met eligibility criteria (see Table A1 in online supplemental material for included studies and their classification). Fourteen studies (10 unique samples) used the two-topic conflict interaction paradigm, and only one (Jayamaha, Antonellis, & Overall, 2016 [Study 3]) of the unique samples ensured relative importance. Thirty-four (26 unique) used a single conflict interac-

tion, and only one (Weldon, Schermerhorn, & Stickle, 2019) ensured the topic was of equal dissatisfaction to both partners. This review intended only to identify how often relative importance was considered, not whether study conclusions rely on these conditions, so this is not a critique of these studies. Instead, these results indicate that relative importance (and its analog in single-conflict studies) has been overlooked by most of the existing literature.

### Study 2

Using two empirical samples, Study 2 estimated how often relative importance is violated when it is not ensured and how often topics in a single-conflict study are rated unequally by partners. Study 2 was conducted in an exploratory manner.

### Method

**Participants.** Sample 1 includes 134 married, severely and stably distressed heterosexual couples from a randomized clinical trial of behavioral couple therapies Christensen et al. (2004). Participants self-identified as Caucasian (77.6%), African American (7.5%), Latinx (5.2%), Asian/Pacific Islander (5.3%), or Native American/Alaskan (0.4%). Mean age was 43.5 years (*SD* = 8.7) for men and 41.6 years (*SD* = 8.6) for women. Mean relationship satisfaction on the Dyadic Adjustment Scale (Spanier, 1976) was 84.5 (*SD* = 15.0) for men and 84.7 (*SD* = 14.0) for women, below the clinical distress cutoff of 97.5.

Sample 2 includes 48 cohabiting community couples who were either monogamously dating for at least one year or married for less than two years. Recruitment was stratified to obtain a range of relationship functioning (for details, see Crenshaw et al., 2019). Participants self-identified as White (76%), Asian (11.5%), Native Hawaiian/Pacific Islander (2.1%), American Indian or Alaska Native (1%), or declined to answer (9.4%). An additional 11 participants (11.5%) identified as Hispanic. Forty-six couples were mixed-gender and two were same gender. Mean age was 26.2

<sup>2</sup> The literature search was also performed without restriction to conflict interactions. Those interested may contact the first author for this list of studies.

years ( $SD = 4.2$ ) for men and 24.7 years ( $SD = 3.5$ ) for women, and mean relationship satisfaction was 67.3 ( $SD = 10.6$ ) for men and 64.2 ( $SD = 17.5$ ) for women on the CSI-16 (Funk & Rogge, 2007), above the clinical distress cutoff of 51.5.

**Procedure.** In Sample 1, couples completed a 2 to 3-hr laboratory session prior to beginning therapy, and we focus on two conflict discussions from this session. To prepare for the discussions, each partner individually rated their dissatisfaction with aspects of the relationship on the Problem Areas Questionnaire (PAQ; Heavey, Christensen, & Malamuth, 1995). Experimenters selected each partner's highest rated issue to be their issue to discuss in one of two, 10-min discussions. If both partners rated the same topic the highest, the topic was selected for the person whose conversation was first. Order of conflict discussions was randomized and counterbalanced.

In Sample 2, couples completed a 3-hr laboratory session that included one 10-min conflict discussion. To choose the topic, each partner completed the PAQ and the experimenter chose the topic with the highest cumulative dissatisfaction across both partners. If two or more topics were tied, the topic more equally rated across partners was selected. Any remaining ties were resolved by asking the couple to choose among the tied topics. All procedures were IRB approved, at the University of California, Los Angeles and the University of Washington (Sample 1) and the University of Utah (Sample 2).

**Measures.** All participants rated their dissatisfaction about 21 common areas of disagreement in romantic relationships (plus two optional write-in topics) on a 7-item Likert scale, from "completely dissatisfied" to "completely satisfied" using the Problem Areas Questionnaire (PAQ; Heavey et al., 1995). Higher scores represent more dissatisfaction.

## Results and Discussion

In Sample 1, relative importance was considered met if the topic "chooser" (the person for whom a topic is selected)<sup>3</sup> was more dissatisfied about the topic than the partner, slightly violated if chooser and partner were equally dissatisfied, or severely violated if the chooser was less dissatisfied. Although Sample 2 included only one conflict discussion, we emulated the selection of two topics in this sample in a way highly similar to Sample 1 to obtain a second estimate of violation frequency (see the [online supplemental material](#) for details and see <https://osf.io/gdp3e/> for full annotated Stata code for this procedure). Finally, we compared ratings for the single topic that was really discussed in Sample 2 to determine how often it was unequally rated.

In Sample 1, when participants discussed a write-in topic (6% of conversations), partner ratings were missing, preventing classification. Shown [Table A2](#) of the online supplemental material, 167 (66%) of the remaining conversations met relative importance, 48 (19%) slightly violated it, and 38 (15%) severely violated it. In Sample 2, 66 (69%) met relative importance, 18 (19%) slightly violated it, and 12 (13%) severely violated it. Violations did not significantly differ by topic in Sample 1 (Odds Ratio [OR] = 1.19,  $p = .509$ ), but were significantly more likely in men's (vs. women's) topics in Sample 2 (OR = 5.08,  $p = .008$ ). In Sample 2, for the single conversation, partners were equally dissatisfied about the topic in 10 (21%) couples, the woman was more dissatisfied in 26 (54%) couples, the man was more dissatisfied in 10 (21%)

couples, and the two same-gender couples (4%) rated the topics differently but could not be compared by gender.

In summary, results were consistent across both samples in finding that relative importance was violated in one third of conversations when not ensured, and there was mixed evidence that violations were more likely in men's topics. For single conflict discussions, female partners were typically more dissatisfied with the topics. Considered in the context of results from Study 1, these findings suggest it is probable that the relative importance assumption is violated in a large proportion of the literature. They also suggest that, often despite researcher intentions, jointly selected single-topic conflict discussions are typically of greater dissatisfaction to the woman in mixed-gender couples. This pattern is consistent with past research that finds women typically desire more change in mixed-gender relationships (e.g., Heyman et al., 2009).

## Study 3

### Method

Study 3 tested whether relative importance assumption violations observed in Study 2 impact results for within-couple tests of behavior, focusing on gender differences in D/W, and clarified previously inconsistent results for three theories of D/W. D/W is a behavior pattern in which one member of a couple nags, complains, criticizes, or otherwise attempts to create change in the other partner, while the other partner avoids or withdraws from the interaction (Christensen & Heavey, 1990). Higher levels of D/W are associated with myriad negative relationship and individual outcomes, including poor marital adjustment and satisfaction, divorce, infidelity, intimate partner violence, depression, anxiety, and alcoholism (Schrodt, Witt, & Shimkowski, 2014). Moreover, D/W plays a crucial role in theoretical models of the course of relationship distress and is often a principal target in couple therapy (e.g., Christensen et al., *in press*).

Three classes of theories have been proposed to explain variability in D/W behavior. *Individual differences* theories suggest that stable differences between men and women drive D/W behavior (Heavey et al., 1993), either due to gender role socialization or differences in physiological reactivity to conflict (see Holley, Sturm, & Levenson, 2010), leading women to engage in demanding behaviors to nurture intimacy or improve the relationship and men to withdraw to maintain independence or avoid conflict. The *conflict structure* theory argues that D/W is determined by the partner who most wants change (the demander) being dependent on the other partner (the withdrawer) to grant the desired change (Heavey et al., 1993). This theory argues that D/W roles and behaviors change fluidly depending on the topic, and roles can switch entirely from one topic to the next. Finally, *social structure* theory is similar to conflict structure in that interest in the particular topic drives D/W behavior. However, social structure argues that, due to power differentials in society, men are abler than

<sup>3</sup> We use this terminology for ease of communication. However, in line with Heyman's (2001) recommendations, the experimenter in these studies ultimately selected the topics based on questionnaire responses, rather than the couple.



women to structure relationships to their wishes (see Holley et al., 2010). Men therefore tend to be more satisfied with the status quo and seek to preserve it (e.g., by withdrawing from change conversations), whereas women tend to be less satisfied with the status quo and seek to change it (e.g., by making requests or demands).

Outlined in Figure 2, the three theories imply specific predictions for how D/W behavior is related to gender and whose topic is being discussed. Individual differences predicts that D/W should be stable regardless of the topic (Figure 2, panel A; Heavey et al., 1993). Conflict structure and social structure predict that D/W should vary by the specific topic under discussion, with D/W roles switching entirely based on the topic in conflict structure (“cross” pattern; Figure 2, panel B), and only partially in social structure (“pinch” pattern; Figure 2, panel C). Integral to conflict structure is that each person demands more and withdraws less than the partner during their own topic, known as *D/W Polarization*. A combination of conflict structure and social structure is a “cross+pinch” pattern, in which D/W roles change by topic, but the change is less pronounced in the man’s topic (Figure 2, panel D). Several studies testing these theories with the two-topic interaction paradigm find that D/W indeed changes across topics, such that partners demand more and withdraw less during their own topic compared with the partner’s topic and significant D/W polarization occurs in the woman’s topic (woman-demand/man-withdraw > man-demand/woman-withdraw). However, only some studies find significant D/W polarization in the man’s topic (Christensen & Heavey, 1990; Holtzworth-Munroe, Smutzler, & Stuart, 1998; Klinetob & Smith, 1996), while others do not (B. Baucom, McFarland, & Christensen, 2010; Heavey et al., 1993).

One significant limitation of this research relates to the relative importance assumption. When testing the effect of gender, topic, and their interaction on D/W behavior, evidence for individual differences rests on the gender main effect absent a topic by gender interaction. Evidence for conflict structure rests on the interaction and the presence of D/W polarization during both partner’s topics. Finally, evidence for social structure rests on the interaction and the presence of polarization during the woman’s topic, but not

during the man’s topic. However, topics become indistinguishable when the relative importance assumption is not met, causing gender and topic effects to be confounded. Consequently, conversations that violate relative importance cannot be used to test these theories, and their inclusion is likely to bias results.

Study 3 tested the effect of relative importance violations on the associations between gender, topic, and observed D/W behavior to identify if and to what extent these violations bias theoretical tests. We predict that the three key tests of conflict structure—the topic\*gender interaction and D/W polarization during each topic—will be significantly smaller when the assumption is violated. We also predict that, consistent with conflict- and social structure theories, greater dissatisfaction discrepancies will be associated with more D/W polarization.

**Participants and procedure.** See Study 2.

#### Measures.

**Demand/withdraw behavior (D/W).** In Sample 1, we measured D/W behavior using the Couples Interaction Rating System-2 (CIRS2; Heavey, Gill, & Christensen, 2002). Trained coders watched each interaction and rated each partner on 13 behaviors, from 1 (*none*) to 9 (*a lot*). Consistent with the CIRS2, we averaged two items—*blame* and *pressure for change*—for the *demand* scale, and three items—*withdraws*, *avoidance*, and reverse-coded *discussion*—for the *withdraw* scale, adding one partner’s *demand* scale to the other’s *withdraw* scale to compute D/W (e.g., man demand/woman withdraw = man’s demand + woman’s withdraw). Equipment failure resulted in loss of D/W data for one couple. Interrater Cronbach’s alphas were .84 to .91 for demand and .82 to .85 for withdraw, and scale alphas were .78 to .84 for demand and .64 to .83 for withdraw.

In Sample 2, we measured D/W using the Asymmetric Behavior Coding System (ABCS; Leo, Crenshaw, & Baucom, 2016). Trained coders rated 22 behaviors on a scale from 1 (absence of behavior) to 7 (extreme form of behavior). For consistency with the CIRS2, we limited demand items to the average of *blame* and *pressure for change* and withdraw items to the average of *withdraws* and *avoidance*, which were worded nearly identically to the

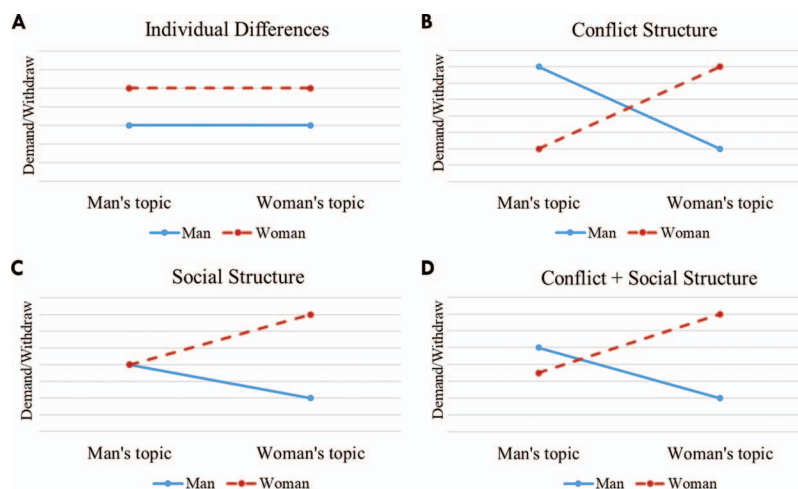


Figure 2. Predicted patterns in demand/withdraw or dissatisfaction ratings according to theories of demand/withdraw behavior. Gender refers to the gender of the demander (e.g., gender = man for man-demand/woman-withdraw). See the online article for the color version of this figure.

CIRS2 items (internal reliability data suggests ABCS coders interpreted the *discussion* item differently from CIRS2 coders, so this item is omitted from the withdraw scale). Results using the full ABCS D/W scales are in the [online supplemental material](#). Interrater alphas were .87 for demand and .89 for withdraw, and scale alphas were .80 for demand and .84 for withdraw.

**Topic dissatisfaction ratings.** See Study 2.

**Analytic plan.** All analyses were conducted in Stata 15 (Stata-Corp, 2017) using multilevel models with restricted maximum likelihood estimation, where partners and conversations are both nested within couples. Consistent with past work (e.g., Christensen & Heavey, 1990), we regressed D/W onto the gender of the demander for each behavior (e.g., male for man-demand/woman-withdraw), topic (man's or woman's topic), and their interaction, using the following model: (presented in series of equations format):

$$\text{Level 1: } D/W_{ij} = \beta_{0j} + \beta_{1j} \times (\text{GENDER}_{ij}) + \beta_{2j} \times (\text{TOPIC}_{ij}) + \beta_{3j} \times (\text{TOPIC} \times \text{GENDER}_{ij}) + r_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \mu_{0j}; \text{ for } i = 1 \text{ to } 3, \beta_{ij} = \gamma_{i0}$$

where  $i$  represents partners within a couple,  $j$  represents couples, GENDER of the demander is effect-coded (*male* =  $-.5$ ; *female* =  $.5$ ), and TOPIC is effect-coded (*male topic* =  $-.5$ ; *female topic* =  $.5$ ). To test if results varied according to relative importance assumption violations, we added dummy-coded assumption status (*assumption met* = 0; *assumption violated* = 1) as an interaction with all terms. We estimated this model to have at least .8 power to detect effect sizes of  $d = .31$  or greater.<sup>4</sup> Lastly, we used single predictor regression to test whether differences in dissatisfaction predicted D/W polarization, which we estimated to have at least .8 power to detect effect sizes of  $r = .175$  for Sample 1 and  $r = .382$  for Sample 2.

## Results

**Planned analyses.** Table A3 displays full results of the model regressing D/W behavior onto the 3-way interaction between gender, topic, and whether the relative importance assumption was met for that conversation, plus all lower order terms. As predicted, the 3-way interaction was significant ( $B = -2.92$ , 95% CI  $[-4.22, -1.62]$ ,  $p < .001$ ), such that the topic\*gender interaction was significantly smaller and was not significantly different from zero when the assumption was violated (Met:  $B = 3.74$ ,  $p < .001$ ; Violated:  $B = .82$ ,  $p = .129$ ; see [Figure 3](#)). These results suggest that this empirical test for conflict structure is susceptible to relative importance assumption violations.

Also consistent with predictions, there was a significant gender\*assumption interaction in the man's topic ( $B = 1.55$ , 95% CI  $[.65, 2.45]$ ,  $p = .001$ ), indicating that D/W polarization during the man's topic was significantly smaller and not significantly different from zero when the assumption was violated (Met:  $B = -.91$ ,  $p = .001$ ; Violated:  $B = .64$ ,  $p = .080$ ). This result suggests that evidence for D/W polarization during the man's topic is impacted by relative importance assumption violations. There was also a significant gender\*assumption interaction in the woman's topic ( $B = -1.37$ , 95% CI  $[-2.31, -.43]$ ,  $p = .004$ ), indicating that D/W polarization during the women's topics was also significantly smaller, but nonetheless still present, when the assumption was

violated (Met:  $B = 2.83$ ,  $p < .001$ ; Violated:  $B = 1.46$ ,  $p < .001$ ). However, even for conversations meeting the relative importance assumption, there was significantly more D/W polarization during the woman's topic compared with the man's topic,  $\chi^2[1] = 24.6$ ,  $p < .001$ , consistent with conflict structure and social structure theories.

Lastly, as predicted, larger discrepancies between topic chooser and partner dissatisfaction ratings were associated with significantly larger D/W polarization (Sample 1:  $B = .41$ , 95% CI  $[.21, .60]$ ,  $p < .001$ ; Sample 2:  $B = .24$ , 95% CI  $[.01, .47]$ ,  $p = .039$ ). These results replicate past work that differences in feelings about the topic predict D/W polarization (e.g., Holley et al., 2010), are consistent with conflict structure and social structure, and suggest that gender differences in D/W in studies using a single conflict discussion may be confounded by differences in satisfaction with the topic.

**Follow-up: Are there negative consequences to ensuring relative importance is met?** While the above results show that neglecting relative importance can have negative consequences when testing within-couple effects, imposing additional restrictions on topic selection may itself have undesirable side effects. To test for these, we emulated topic selection while requiring the assumption be met (see [online supplemental material](#) for procedure details), addressing four questions: 1) How often is it impossible to meet the assumption? 2) Does requiring the assumption result in less important topics? 3) Does the discrepancy between chooser and partner ratings change? 4) Which scenario (requiring the assumption vs. not) results in topics that are more equivalent across gender?

Nine conversations (of 268; 3.4%) in Sample 1 and three (of 96; 3.1%) in Sample 2 could not meet relative importance because one partner was equally or more dissatisfied about every topic. All such cases could, however, meet the less strict criterion of allowing an equally rated topic. When the assumption was violated, selecting a new topic that met the assumption did not result in a significantly lower-rated topic for the chooser in Sample 1 ( $B = -0.1$ ,  $p = .725$ ), but resulted in a topic rated 0.6 points (on a 1–7 scale) lower on average in Sample 2 ( $p < .001$ ).

In line with the goal of selecting clearly distinguishable topics, the discrepancy between chooser and partner ratings was between 1.4 (Sample 2) and 2.9 (Sample 1) points higher when topics that violated the assumption were replaced with ones that did not ( $ps < .001$ ). To test which procedure (requiring the assumption vs. not) resulted in topics that functioned most similarly across genders, we tested for associations between gender and chooser dissatisfaction, partner dissatisfaction, and chooser-partner discrepancy in the full samples both when the assumption was required and when it was not. These analyses, the full details of which are in the [online](#)

<sup>4</sup> For tests of the effects of assumption violations, we estimated power by adjusting the effective sample size to account for data dependence in multilevel models using the formula  $N_{\text{effective}} = \frac{N}{1 + [n - 1] * \rho}$ , in which  $N$  is the total number of observations (532),  $n$  is the number of observations within the couple (4), and  $\rho$  is the level 2 intraclass correlation coefficient (ICC; .195; Diggle et al., 2002). This adjustment resulted in  $N_{\text{effective}} = 335$ . This adjustment was not needed for tests of the effect of differences in dissatisfaction on D/W polarization because the ICC for Sample 1 was 0 (i.e., no group-level dependence), and Sample 2 only had 1 observation per couple.

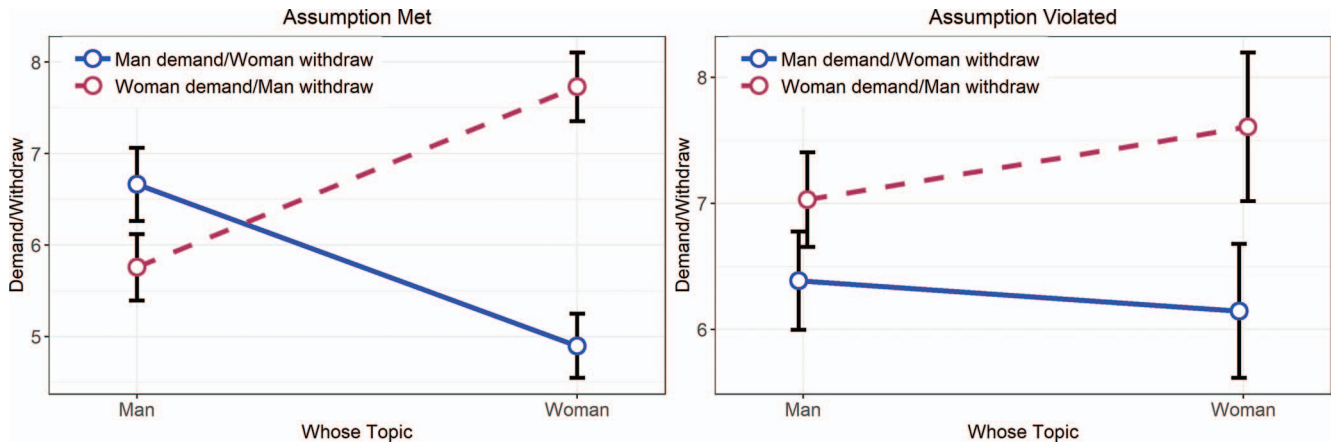


Figure 3. Associations between topic and gender and observed D/W behavior when the relative importance assumption is met (left panel) and when it is violated (right panel). Error bars represent 95% confidence intervals of the mean. See the online article for the color version of this figure.

supplemental material, showed that the two procedures both produced similarly rated topics across genders, but found that requiring the assumption performed marginally better for minimizing systematic gender differences in topic ratings.

## Discussion

Study 3 found that associations between D/W behavior, topic, and gender significantly varied according to whether or not the relative importance assumption was met, including all three tests of conflict structure, indicating that statistical tests of such associations may be confounded by assumption violations in studies that do not ensure the assumption is met. We also found that gender differences in single-topic studies of D/W behavior may be confounded by satisfaction with the topic. Requiring the assumption had minimal negative side effects, and there was marginal evidence it produced slightly greater equivalency of topics across gender. These results support the importance of attending to the relative importance assumption.

## General Discussion

Behavioral observation of how couples navigate conflict constitutes a major focus of research on intimate relationships, research that informs clinical practice. The present study introduced the *relative importance assumption*—the assumption that each partner is more concerned or dissatisfied with their topic than the partner is—for tests of within-couple effects in the conflict interaction paradigm. The relative importance assumption must be met in order for the topics to be distinguishable from one another in studies where each partner selects a topic for discussion, a necessary condition for drawing clear conclusions about topic effects, gender differences in mixed-gender couples, and other within-couple effects.

The present investigation found that only one of 10 studies using the two-topic paradigm ensured relative importance was met, and it was violated in one third of conversations when not ensured, suggesting this issue is both overlooked in much of the literature and is likely to be violated frequently. We also tested the practical

consequences of relative importance assumption violations for tests of within-couple effects in gender differences in D/W and found strong evidence that these associations are impacted by assumption violations. All three key tests of the conflict structure theory varied significantly—and two of three tests were rendered nonsignificant—when the assumption was violated. These results suggest that relative importance assumption violations can obfuscate theoretical tests and may lead to incorrect conclusions. We also found that a common procedure for selecting a single topic resulted in a topic rated more highly by female partners a majority of the time, lending support to Heavey et al.'s (1993) supposition that single-conflict studies “have probably disproportionately observed interactions in which the wife is the person primarily invested in the discussion” (p. 26).

These results may hold broader implications for the state of knowledge about the interactions of romantic partners and its application to clinical practice. For one, research with couples in which female partners are overall more dissatisfied with the topic may overestimate gender differences in behavior during relationship interactions. Similarly, couple therapists may overdiagnose gender-stereotyped behavior or miss other important behavior patterns if issues discussed in therapy are decided upon jointly by the couple, or when one partner is more dissatisfied with both individually raised issues. Even if reflecting actual realities of mixed-gender relationships—that relationships generally create greater satisfaction for men than women (e.g., Jacobson, 1983)—researchers and clinicians might well attribute findings to gendered differences in behavior rather than to gendered differences in the structure of the relationship.

## Drivers of Demand/Withdraw Behavior

By restricting analyses only to those conversations meeting the relative importance assumption, the present studies could test theories about the drivers of D/W behavior more precisely than past research and clarify previously inconsistent results. D/W roles changed depending on whose topic was being discussed, wherein partners were more dissatisfied, demanded more, and withdrew less during their own topic compared with their partner and com-

pared with their own behavior during the partner's topic. However, we also observed greater D/W polarization during women's topics compared with men's, resulting in the "cross+pinch" pattern consistent with both conflict structure and social structure theories. Additionally, D/W polarization during the man's topic significantly varied by assumption status and was observed for conversations meeting the assumption, but not those that did not meet the assumption.

We take these results to support the conflict structure and social structure theories of D/W. Specifically, conflict structure theory appears differentially applicable to men and women in mixed-gender relationships: polarization was only present in men's topics that met the relative importance assumption while it was broadly present in women's topics. These findings suggest women tend to take on a demanding role in topics they deem important regardless of whether or not the partner is similarly dissatisfied. However, men only take on a demanding role when the topic is more important to them than to their female partner. These findings are consistent with a wide body of literature detailing women's tendency to do more emotional work, identify more areas of discontent in relationships, and engage in more relationally schematic processing than men, and men's tendency to selectively participate in a range of relationship maintenance behaviors (e.g., house work, parenting; Jacobson, 1983). It would be valuable for future research to explore why women and men tend to rate relationship problems in this way (e.g., motivation, or attributions of responsibility to address the problem) and to identify factors that may attenuate these differences (e.g., egalitarian gender role identification; Markman, Silvern, Clements, & Kraft-Hanak, 1993).

Lastly, these results help explain past inconsistency regarding D/W polarization during men's topics. Studies that do not find significant D/W polarization during men's topic (B. Baucom et al., 2010; Heavey et al., 1993) may have a greater number of assumption violations than those that do, masking the effect. It would be valuable for future work to report the percentage of assumption violations so that this possibility can be further explored.

## Limitations

There are several limitations to the current studies. First, topic-selection procedures for the two samples were highly similar to one another and are not necessarily representative of the broader literature. However, the current samples employed careful topic-selection procedures in line with Heyman's (2001) recommendations, in contrast to many studies in the literature review, so this difference may actually result in *underestimating* the rate of assumption violations. Second, it is possible that some studies ensured relative importance was met but did not report it, which we could confirm in two instances by Overall and colleagues (Overall, Fletcher, Simpson, & Fillo, 2015; Overall, Simpson, & Struthers, 2013), which did not report using procedures ensuring relative importance, but used the same sample as a study that did (Jayamaha et al., 2016). It is possible there are other such cases.

Third, although we do not have reason to believe it is the case, men and women may rate their dissatisfaction on PAQ items in different ways. Similarly, partners' overall feelings about the relationship could have colored how they feel about individual topics, causing topic ratings to be pulled toward the overall relationship sentiment. Ample evidence from the current and past

investigations demonstrate that feelings about particular issues predict behavior in expected ways (e.g., Holley et al., 2010), but we cannot be sure that equivalent ratings on PAQ topics are truly indicative of equivalence on the underlying construct. Finally, although the relative importance assumption still applies to interactions of same gender couples, lack of representation in our samples limits generalizability to these couples.

## Recommendations and Future Directions

We provide several recommendations for choosing topics in observational research, focusing on studies of conflict, and suggest reporting standards to make key procedural details explicit and to facilitate interpretation when results differ across studies. We first reiterate Heyman's (2001) four recommendations that researchers (a) select the topics themselves, (b) narrow down broad topics via questionnaire or play-by-play interview, (c) standardize and report task instructions to couples, and (d) control the gender of the topic chooser by choosing two topics (one for each partner) or keeping the chooser's gender constant. Ensuring relative importance would require one additional step: after tentatively selecting topics through whichever criteria are used (e.g., each partner's highest-rated topic), examine partner ratings for those topics to ensure the partner did not rate the topic higher; if they did, move on to the next topic.

The decision of whether to ensure relative importance depends on a variety of factors. As we have outlined, it is necessary when the theoretical question relies on distinguishable topics, such as tests of topic effects (e.g., whose topic is discussed or partner roles in topics) and tests of variables that differ within a couple (e.g., gender differences in mixed-gender couples). In these cases, not meeting the assumption can result in irreparable and fatal confounding of topic with other effects. In contrast, enforcing relative importance is *not* necessary when the aim is primarily descriptive or sociological. For example, a study making descriptive comparisons to draw conclusions about how couples tend to behave across cultures (e.g., Tsai, Levenson, & McCoy, 2006) may want to exert minimal control over topic selection so interactions are as ecologically valid as possible. Such studies cannot, however, speak to within-couple causes or drivers of observed behavior without ensuring relative importance of topics.

Other cases, such as behavior change over time and tests of between-couple variables, are less straightforward. When in doubt about whether to require relative importance, a rule of thumb is to consider whether a study's conclusions would change if one partner chose *both* topics, as this situation parallels the practical consequence of violating relative importance. Using this logic, we expect relative importance to be less important for these cases. However, not ensuring relative importance, or at least controlling for it, assumes it is not relevant. In these cases—and in general—it may be beneficial to articulate and test such assumptions. For example, showing that relative importance is unrelated to the behavior of interest would provide justification for not ensuring it. Additionally, if relative importance is not consequential, then it should not matter much whose topic is discussed; therefore, behavior should be highly correlated across topics. If behavior significantly varies across topics, that would suggest topic may be more influential than expected and may need to be accounted for. Testing assumptions can also allow for exceptions to the general



rule of ensuring relative importance for tests of within-couple variables: if neither topic nor its interaction with the focal predictor are associated with the behavior of interest, then ensuring relative importance may not be necessary.

Fortunately, it is often possible to maximize multiple goals at once with minimal added procedural complication, which we recommend in general. For example, ensuring partners always discuss their most highly rated topics *while meeting relative importance as often as possible* requires just two steps: (a) select the highest-rated topic, and (b) if multiple topics are rated the highest (a common situation), select one that meets or is as close as possible to meeting relative importance. Conversely, if meeting relative importance is the primary goal: select the highest-rated topic only from among topics meeting the assumption. Finally, when meeting relative importance is necessary or desired, but impossible because one partner rates all topics equally or higher than the other, experimenters can select the topic that is closest to meeting it. However, for tests that rely on topic distinguishability, we recommend conducting sensitivity analyses with and without these couples to determine whether including them changes results.

## Conclusion

The present research demonstrated that observational investigations of conflict in intimate relationships (a) often choose topics that are more of interest for one partner than the other, and (b) such choices can impact the results of those studies. Relationship researchers and couple therapists may miss or overestimate certain behaviors when observing interactions without placing appropriate constraints. We recommend investigators carefully choose discussion topics for observational research and couple therapy based on the underlying goals of the study, articulate and test assumptions of behavioral tasks, and, unless it can be demonstrated to be unimportant, ensure the relative importance assumption for tests of within-couple variables.

## References

References marked with an asterisk (\*) indicate studies included in the literature review.

- \*Allred, K. M., Chambless, D. L., Porter, E., Brier, M. J., & Schwartz, R. A. (2018). Attributions and perceptions of criticism: An examination of patients with anxiety and normal control participants. *Journal of Family Psychology, 32*, 947–956. <http://dx.doi.org/10.1037/fam0000460>
- \*Bakhurst, M., McGuire, A., & Halford, W. K. (2018). Trauma Symptoms, communication, and relationship satisfaction in military couples. *Family Process, 57*, 241–252. <http://dx.doi.org/10.1111/famp.12285>
- \*Barry, R. A., & Lawrence, E. (2013). “Don’t stand so close to me”: An attachment perspective of disengagement and avoidance in marriage. *Journal of Family Psychology, 27*, 484–494. <http://dx.doi.org/10.1037/a0032867>
- \*Baucom, B. R., Atkins, D. C., Rowe, L. S., Doss, B. D., & Christensen, A. (2015). Prediction of treatment response at 5-year follow-up in a randomized clinical trial of behaviorally based couple therapies. *Journal of Consulting and Clinical Psychology, 83*, 103–114. <http://dx.doi.org/10.1037/a0038005>
- \*Baucom, B. R., Dickenson, J. A., Atkins, D. C., Baucom, D. H., Fischer, M. S., Weusthoff, S., . . . Zimmermann, T. (2015). The interpersonal process model of demand/withdraw behavior. *Journal of Family Psychology, 29*, 80–90. <http://dx.doi.org/10.1037/fam0000044>
- Baucom, B. R., McFarland, P. T., & Christensen, A. (2010). Gender, topic, and time in observed demand-withdraw interaction in cross- and same-sex couples. *Journal of Family Psychology, 24*, 233–242. <http://dx.doi.org/10.1037/a0019717>
- \*Beck, L. A., Pietromonaco, P. R., DeBuse, C. J., Powers, S. I., & Sayer, A. G. (2013). Spouses’ attachment pairings predict neuroendocrine, behavioral, and psychological responses to marital conflict. *Journal of Personality and Social Psychology, 105*, 388–424. <http://dx.doi.org/10.1037/a0033056>
- \*Birnbaum, G. E., Mikulincer, M., & Austerlitz, M. (2013). A fiery conflict: Attachment orientations and the effects of relational conflict on sexual motivation. *Personal Relationships, 20*, 294–310. <http://dx.doi.org/10.1111/j.1475-6811.2012.01413.x>
- \*Campbell, L., Overall, N. C., Rubin, H., & Lackenbauer, S. D. (2013). Inferring a partner’s ideal discrepancies: Accuracy, projection, and the communicative role of interpersonal behavior. *Journal of Personality and Social Psychology, 105*, 217–233. <http://dx.doi.org/10.1037/a0033009>
- \*Cao, H., Fang, X., Fine, M. A., Ju, X., Lan, J., & Liu, X. (2015). Beyond the average marital communication: Latent profiles of the observed interactions among Chinese newlywed couples. *Journal of Family Psychology, 29*, 850–862. <http://dx.doi.org/10.1037/fam0000126>
- Christensen, A., Atkins, D. C., Berns, S., Wheeler, J., Baucom, D. H., & Simpson, L. E. (2004). Traditional versus integrative behavioral couple therapy for significantly and chronically distressed married couples. *Journal of Consulting and Clinical Psychology, 72*, 176–191. <http://dx.doi.org/10.1037/0022-006X.72.2.176>
- Christensen, A., Doss, B. D., & Jacobson, N. S. (in press). *Integrative behavioral couple therapy: A therapist’s guide to creating acceptance and change*. New York, NY: Norton.
- Christensen, A., & Heavey, C. L. (1990). Gender and social structure in the demand/withdraw pattern of marital conflict. *Journal of Personality and Social Psychology, 59*, 73–81. <http://dx.doi.org/10.1037/0022-3514.59.1.73>
- Cohen, O., & Levin, L. (2012). Impact of financial and employment status on the co-parenting of divorcing couples in Israel. *Journal of Family Studies, 18*, 47–61. <http://dx.doi.org/10.5172/jfs.2012.18.1.47>
- Crenshaw, A. O., Leo, K., & Baucom, B. R. W. (2019). The effect of stress on empathic accuracy in romantic couples. *Journal of Family Psychology, 33*, 327–337. <http://dx.doi.org/10.1037/fam0000508>
- Diggle, P., Diggle, P. J., Heagerty, P., Heagerty, P. J., Liang, K.-Y., & Zeger, S. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford, UK: Oxford University Press.
- \*Ellison, J. K., Kouros, C. D., Papp, L. M., & Cummings, E. M. (2016). Interplay between marital attributions and conflict behavior in predicting depressive symptoms. *Journal of Family Psychology, 30*, 286–295. <http://dx.doi.org/10.1037/fam0000181>
- Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology, 21*, 572–583. <http://dx.doi.org/10.1037/0893-3200.21.4.572>
- \*Girme, Y. U., Overall, N. C., Simpson, J. A., & Fletcher, G. J. O. (2015). “All or nothing”: Attachment avoidance and the curvilinear effects of partner support. *Journal of Personality and Social Psychology, 108*, 450–475. <http://dx.doi.org/10.1037/a0038866>
- \*Gordon, A. M., & Chen, S. (2016). Do you get where I’m coming from? Perceived understanding buffers against the negative impact of conflict on relationship satisfaction. *Journal of Personality and Social Psychology, 110*, 239–260. <http://dx.doi.org/10.1037/pspi0000039>
- \*Hammett, J. F., Karney, B. R., & Bradbury, T. N. (2018). Longitudinal effects of increases and decreases in intimate partner aggression. *Journal of Family Psychology, 32*, 343–354. <http://dx.doi.org/10.1037/fam0000351>

- \*Hanley, K. E., Leifker, F. R., Blandon, A. Y., & Marshall, A. D. (2013). Gender differences in the impact of posttraumatic stress disorder symptoms on community couples' intimacy behaviors. *Journal of Family Psychology, 27*, 525–530. <http://dx.doi.org/10.1037/a0032890>
- Heavey, C. L., Christensen, A., & Malamuth, N. M. (1995). The longitudinal impact of demand and withdrawal during marital conflict. *Journal of Consulting and Clinical Psychology, 63*, 797–801. <http://dx.doi.org/10.1037/0022-006X.63.5.797>
- Heavey, C., Gill, D., & Christensen, A. (2002). *The Couples Interaction Rating System-2*. Unpublished rating manual, University of California, Los Angeles, CA.
- Heavey, C. L., Layne, C., & Christensen, A. (1993). Gender and conflict structure in marital interaction: A replication and extension. *Journal of Consulting and Clinical Psychology, 61*, 16–27. <http://dx.doi.org/10.1037/0022-006X.61.1.16>
- Heyman, R. E. (2001). Observation of couple conflicts: Clinical assessment applications, stubborn truths, and shaky foundations. *Psychological Assessment, 13*, 5–35.
- Heyman, R. E., Hunt-Martorano, A. N., Malik, J., & Slep, A. M. S. (2009). Desired change in couples: Gender differences and effects on communication. *Journal of Family Psychology, 23*, 474–484. <http://dx.doi.org/10.1037/a0015980>
- \*Hiew, D. N., Halford, W. K., van de Vijver, F. J. R., & Liu, S. (2016). Communication and relationship satisfaction in Chinese, Western, and intercultural Chinese-Western couples. *Journal of Family Psychology, 30*, 193–202. <http://dx.doi.org/10.1037/fam0000144>
- Holley, S. R., Sturm, V. E., & Levenson, R. W. (2010). Exploring the basis for gender differences in the demand-withdraw pattern. *Journal of Homosexuality, 57*, 666–684. <http://dx.doi.org/10.1080/00918361003712145>
- Holtzworth-Munroe, A., Smutzler, N., & Stuart, G. L. (1998). Demand and withdraw communication among couples experiencing husband violence. *Journal of Consulting and Clinical Psychology, 66*, 731–743. <http://dx.doi.org/10.1037/0022-006X.66.5.731>
- Jacobson, N. S. (1983). Beyond empiricism: The politics of marital therapy. *American Journal of Family Therapy, 11*, 11–24. <http://dx.doi.org/10.1080/01926188308250118>
- \*Jayamaha, S. D., Antonellis, C., & Overall, N. C. (2016). Attachment insecurity and inducing guilt to produce desired change in romantic partners: Attachment insecurity and partner regulation. *Personal Relationships, 23*, 311–338. <http://dx.doi.org/10.1111/pere.12128>
- \*Kellas, J. K., Carr, K., Horstman, H. K., & Dilillo, D. (2017). The Communicated Perspective-Taking Rating System and links to well-being in marital conflict: Communicated perspective taking. *Personal Relationships, 24*, 185–202. <http://dx.doi.org/10.1111/pere.12177>
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic Data Analysis*. New York, NY: Guilford Press.
- \*Kliem, S., Weusthoff, S., Hahlweg, K., Baucom, K. J. W., & Baucom, B. R. (2015). Predicting long-term risk for relationship dissolution using nonparametric conditional survival trees. *Journal of Family Psychology, 29*, 807–817. <http://dx.doi.org/10.1037/fam0000134>
- Klinetob, N. A., & Smith, D. A. (1996). Demand-withdraw communication in marital interaction: Tests of interspousal contingency and gender role hypotheses. *Journal of Marriage and the Family, 58*, 945–957. <http://dx.doi.org/10.2307/353982>
- \*Kusner, K. G., Mahoney, A., Pargament, K. I., & DeMaris, A. (2014). Sanctification of marriage and spiritual intimacy predicting observed marital interactions across the transition to parenthood. *Journal of Family Psychology, 28*, 604–614. <http://dx.doi.org/10.1037/a0036989>
- \*Kuster, M., Bernecker, K., Backes, S., Brandstätter, V., Nussbeck, F. W., Bradbury, T. N., . . . Bodenmann, G. (2015). Avoidance orientation and the escalation of negative communication in intimate relationships. *Journal of Personality and Social Psychology, 109*, 262–275. <http://dx.doi.org/10.1037/pspi0000025>
- \*Lambert, N., Fincham, F. D., Dewall, N. C., Pond, R., & Beach, S. R. (2013). Shifting toward cooperative tendencies and forgiveness: How partner-focused prayer transforms motivation. *Personal Relationships, 20*, 184–197. <http://dx.doi.org/10.1111/j.1475-6811.2012.01411.x>
- \*Lee, W.-Y., Nakamura, S., Chung, M. J., Chun, Y. J., Fu, M., Liang, S.-C., & Liu, C.-L. (2013). Asian couples in negotiation: A mixed-method observational study of cultural variations across five Asian regions. *Family Process, 52*, 499–518. <http://dx.doi.org/10.1111/famp.12040>
- \*Lemay, E. P. (2016). The forecast model of relationship commitment. *Journal of Personality and Social Psychology, 111*, 34–52. <http://dx.doi.org/10.1037/pspi0000052>
- \*Lemay, E. P., Jr. (2014). Accuracy and bias in self-perceptions of responsive behavior: Implications for security in romantic relationships. *Journal of Personality and Social Psychology, 107*, 638–656. <http://dx.doi.org/10.1037/a0037298>
- Leo, K., Crenshaw, A. O., & Baucom, B. R. W. (2016). *Asymmetric behavior coding system*. Unpublished manuscript.
- \*Lucas-Thompson, R. G., George, M. W., & Quinn-Sparks, A. R. (2016). Trust: An innovative tool for investigating marital conflict in response to a novel stressor. *Journal of Family Psychology, 30*, 625–632. <http://dx.doi.org/10.1037/fam0000211>
- \*Lucas-Thompson, R. G., & Granger, D. A. (2014). Parent-child relationship quality moderates the link between marital conflict and adolescents' physiological responses to social evaluative threat. *Journal of Family Psychology, 28*, 538–548. <http://dx.doi.org/10.1037/a0037328>
- \*Marigold, D. C., & Anderson, J. E. (2016). Shifting expectations of partners' responsiveness changes outcomes of conflict discussions: Expectations of partners' responsiveness. *Personal Relationships, 23*, 517–535. <http://dx.doi.org/10.1111/pere.12141>
- Markman, H. J., Silvern, L., Clements, M., & Kraft-Hanak, S. (1993). Men and women dealing with conflict in heterosexual relationships. *Journal of Social Issues, 49*, 107–125. <http://dx.doi.org/10.1111/j.1540-4560.1993.tb01171.x>
- \*McHale, J. P., Salman-Engin, S., & Coovert, M. D. (2015). Improvements in unmarried African American parents' rapport, communication, and problem-solving following a prenatal coparenting intervention. *Family Process, 54*, 619–629. <http://dx.doi.org/10.1111/famp.12147>
- \*Neff, L. A., & Geers, A. L. (2013). Optimistic expectations in early marriage: A resource or vulnerability for adaptive relationship functioning? *Journal of Personality and Social Psychology, 105*, 38–60. <http://dx.doi.org/10.1037/a0032600>
- \*Nguyen, T. P., Williamson, H. C., Karney, B. R., & Bradbury, T. N. (2017). Communication moderates effects of residential mobility on relationship quality among ethnically diverse couples. *Journal of Family Psychology, 31*, 753–764. <http://dx.doi.org/10.1037/fam0000324>
- \*Overall, N. C., Fletcher, G. J. O., Simpson, J. A., & Fillo, J. (2015). Attachment insecurity, biased perceptions of romantic partners' negative emotions, and hostile relationship behavior. *Journal of Personality and Social Psychology, 108*, 730–749. <http://dx.doi.org/10.1037/a0038987>
- \*Overall, N. C., Girme, Y. U., Lemay, E. P., Jr., & Hammond, M. D. (2014). Attachment anxiety and reactions to relationship threat: The benefits and costs of inducing guilt in romantic partners. *Journal of Personality and Social Psychology, 106*, 235–256. <http://dx.doi.org/10.1037/a0034371>
- \*Overall, N. C., Hammond, M. D., McNulty, J. K., & Finkel, E. J. (2016). When power shapes interpersonal behavior: Low relationship power predicts men's aggressive responses to low situational power. *Journal of Personality and Social Psychology, 111*, 195–217. <http://dx.doi.org/10.1037/pspi0000059>
- \*Overall, N. C., Simpson, J. A., & Struthers, H. (2013). Buffering attachment-related avoidance: Softening emotional and behavioral defenses during conflict discussions. *Journal of Personality and Social Psychology, 104*, 854–871. <http://dx.doi.org/10.1037/a0031798>

- \*Rauer, A., & Volling, B. (2013). More than one way to be happy: A typology of marital happiness. *Family Process, 52*, 519–534. <http://dx.doi.org/10.1111/famp.12028>
- \*Rauer, A., Williams, L., & Jensen, J. (2017). Finer distinctions: Variability in satisfied older couples' problem-solving behaviors. *Family Process, 56*, 501–517. <http://dx.doi.org/10.1111/famp.12198>
- \*Rentscher, K. E., Rohrbaugh, M. J., Shoham, V., & Mehl, M. R. (2013). Asymmetric partner pronoun use and demand-withdraw interaction in couples coping with health problems. *Journal of Family Psychology, 27*, 691–701. <http://dx.doi.org/10.1037/a0034184>
- \*Ross, J. M., Karney, B. R., Nguyen, T. P., & Bradbury, T. N. (2019). Communication that is maladaptive for middle-class couples is adaptive for socioeconomically disadvantaged couples. *Journal of Personality and Social Psychology, 116*, 582–597. <http://dx.doi.org/10.1037/pspi0000158>
- \*Russell, V. M., Baker, L. R., McNulty, J. K., & Overall, N. C. (2018). "You're forgiven, but don't do it again!" Direct partner regulation buffers the costs of forgiveness. *Journal of Family Psychology, 32*, 435–444. <http://dx.doi.org/10.1037/fam0000409>
- Sagrestano, L. M., Heavey, C. L., & Christensen, A. (1999). Perceived Power and Physical Violence in Marital Conflict. *Journal of Social Issues, 55*, 65–79. <http://dx.doi.org/10.1111/0022-4537.00105>
- Schrodt, P., Witt, P. L., & Shimkowski, J. R. (2014). A meta-analytical review of the demand/withdraw pattern of interaction and its associations with individual, relational, and communicative outcomes. *Communication Monographs, 81*, 28–58. <http://dx.doi.org/10.1080/03637751.2013.813632>
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for Assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38*, 15–28. <http://dx.doi.org/10.2307/350547>
- Stanley, S. M., Bradbury, T. N., & Markman, H. J. (2000). Structural flaws in the bridge from basic research on marriage to interventions for couples. *Journal of Marriage and the Family, 62*, 256–264. <http://dx.doi.org/10.1111/j.1741-3737.2000.00256.x>
- StataCorp. (2017). *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.
- \*Thomson, R. A., Overall, N. C., Cameron, L. D., & Low, R. S. T. (2018). Perceived regard, expressive suppression during conflict, and conflict resolution. *Journal of Family Psychology, 32*, 722–732. <http://dx.doi.org/10.1037/fam0000429>
- \*Thorson, K. R., Lorber, M. F., Slep, A. M. S., & Heyman, R. E. (2018). Adult adiposity linked to relationship hostility for low-cortisol reactors. *Journal of Family Psychology, 32*, 197–205. <http://dx.doi.org/10.1037/fam0000397>
- \*Trumbell, J. M., Hibbel, L. C., Mercado, E., & Posada, G. (2018). The impact of marital withdrawal and secure base script knowledge on mothers' and fathers' parenting. *Journal of Family Psychology, 32*, 699–709. <http://dx.doi.org/10.1037/fam0000402>
- Tsai, J. L., Levenson, R. W., & McCoy, K. (2006). Cultural and temperamental variation in emotional response. *Emotion, 6*, 484–497. <http://dx.doi.org/10.1037/1528-3542.6.3.484>
- \*Van Vleet, M., Helgeson, V. S., Seltman, H. J., Korytkowski, M. T., & Hausmann, L. R. M. (2018). Communally coping with diabetes: An observational investigation using the actor-partner interdependence model. *Journal of Family Psychology, 32*, 654–663. <http://dx.doi.org/10.1037/fam0000414>
- \*Weldon, H. T., Schermerhorn, A. C., & Stickle, T. R. (2019). Children's affective and arousal responses to live interparental conflict: Links with appraisals. *Journal of Family Psychology, 33*, 607–616. <http://dx.doi.org/10.1037/fam0000522>
- \*Weusthoff, S., Baucom, B. R., & Hahlweg, K. (2013). Fundamental frequency during couple conflict: An analysis of physiological, behavioral, and sex-linked information encoded in vocal expression. *Journal of Family Psychology, 27*, 212–220. <http://dx.doi.org/10.1037/a0031887>
- \*Williamson, H. C., Altman, N., Hsueh, J., & Bradbury, T. N. (2016). Effects of relationship education on couple communication and satisfaction: A randomized controlled trial with low-income couples. *Journal of Consulting and Clinical Psychology, 84*, 156–166. <http://dx.doi.org/10.1037/ccp0000056>
- \*Williamson, H. C., Hanna, M. A., Lavner, J. A., Bradbury, T. N., & Karney, B. R. (2013). Discussion topic and observed behavior in couples' problem-solving conversations: Do problem severity and topic choice matter? *Journal of Family Psychology, 27*, 330–335. <http://dx.doi.org/10.1037/a0031534>
- \*Williamson, H. C., Karney, B. R., & Bradbury, T. N. (2013). Financial strain and stressful events predict newlyweds' negative communication independent of relationship satisfaction. *Journal of Family Psychology, 27*, 65–75. <http://dx.doi.org/10.1037/a0031104>
- \*Winer, J. P., Powers, S. I., Pietromonaco, P. R., & Schreck, M. C. (2018). Childhood family adversity and adult cortisol response: The role of observed marital conflict behavior. *Journal of Family Psychology, 32*, 793–803. <http://dx.doi.org/10.1037/fam0000455>
- Woody, E., & Sadler, P. (2005). Structural equation models for interchangeable dyads: Being the same makes a difference. *Psychological Methods, 10*, 139–158. <http://dx.doi.org/10.1037/1082-989X.10.2.139>
- \*Zhou, J., & Davila, J. (2019). Romantic competence behavior during problem solving among emerging adult dating couples: Development of an observational coding system. *Personal Relationships, 26*, 448–465. <http://dx.doi.org/10.1111/pere.12289>

Received September 15, 2019

Revision received June 2, 2020

Accepted June 8, 2020 ■