Published in final edited form as:

Behav Ther. 2011 September ; 42(3): 364–377. doi:10.1016/j.beth.2010.09.003.

A Multilevel Examination of Interpartner Intimate Partner Violence and Psychological Aggression Reporting Concordance

Amy D. Marshall,
The Pennsylvania State University

Jillian Panuzio,
University of Nebraska, Lincoln

Kerry N. Makin-Byrd,
The Pennsylvania State University

Casey T. Taft, and
National Center for PTSD, VA Boston Healthcare System, and Boston University School of Medicine

Amy Holtzworth-Munroe
Indiana University

Abstract

Low concordance of reports across partners has consistently been observed when partners report the frequency of intimate partner violence (IPV) and psychological aggression (PA) in their relationship. Researchers have been unsuccessful in the quest to discover systematic biases across reporters, perhaps due to examining constructs that are not the source of bias (e.g., gender, victim/perpetrator status) or examining potentially fruitful constructs using underpowered statistics or erroneous conceptualizations (e.g., examining variables at a dyadic, rather than an individual, level). We used multilevel modeling with two samples ($Ns = 88$ and 164 couples) to examine husbands' and wives' relationship satisfaction as individual-level correlates of husband and wife-perpetrated IPV and PA reporting concordance. Consistent with prior literature, low to moderate levels of agreement were observed, and gender and victim/perpetrator status were not consistently associated with reporting concordance. In contrast, for both husbands and wives, relationship satisfaction was associated with reporting concordance such that high relationship satisfaction was related to reporting less of one's partner's PA than the partner reported, whereas low relationship satisfaction was related to reporting more of one's partner's PA than the partner reported. A similar pattern of results emerged for the reporting of IPV, but results did not cross validate between samples. These findings suggest that relationship satisfaction may lead to either reluctance, or increased willingness, to attribute negative relationship events to partner behavior, potentially due to partner blame and relationship schemas. In addition, the influence of individual-level factors may be occluded when aggregated across partners to examine correlates of interpartner reporting concordance.

© 2011 Published by Elsevier Ltd on behalf of Association for Behavioral and Cognitive Therapies.

Address correspondence to Amy D. Marshall, Ph.D., Department of Psychology, The Pennsylvania State University, 415 Moore Building, University Park, PA 16802; AmyMarshall@psu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Across clinical and research domains, the collection of behavioral data from multiple reporters is considered a valuable means of increasing assessment validity. However, reporters frequently disagree, even when reporting on relatively objective behaviors. For example, among intimate partners, low concordance of reports has been observed for behaviors such as the frequency of fathers' contact with their children (Mikelson, 2008) and whether the partners engaged in sexual behaviors that day (Jacobson & Moore, 1981). One means of understanding this phenomenon, and ultimately improving assessment methods, is to consider individual characteristics that may differentially influence each reporter's response. In the current study, we examine the concordance of partners' reports of intimate partner violence (IPV) and psychological aggression (PA), and we consider the impact of each partner's relationship satisfaction on his or her reporting behavior. We additionally aim to use this topic area as an example of an advanced statistical method through which researchers can obtain a better understanding of individual-level factors that decrease reliability across multiple reporters.

The issue of interpartner reporting concordance has been a particularly daunting limitation in the assessment of IPV and PA (Armstrong, Wernke, Medina, & Schafer, 2002). Previous investigations of couples in a variety of settings have demonstrated that couples evidence low to moderate levels of agreement regarding the occurrence and frequency of IPV and PA in their relationships (Archer, 1999; Caetano, Field, Ramisetty-Mikler, & Lipsky, 2009). Because IPV and PA primarily occur in private settings without the availability of third-party reporters, and police records underestimate true levels of aggression, without interpartner reliability the validity of such assessments is compromised. Such unreliable reports may lead to inconsistent and/or inaccurate research conclusions, as well as difficulties in the accurate identification of aggression perpetration and victimization in clinical contexts.

Research designed to discern systematic biases that lead to low concordance of IPV and PA reports has generally found that perpetrators report less of their own IPV and PA perpetration than their partners report, but, compared to the reports of their partners, men report less of their own IPV and PA perpetration to a greater extent than do women (Archer, 1999; Perry & Fromuth, 2005). However, several studies have conversely found that men report less IPV and PA than women, regardless of the perpetrator's gender (e.g., Panuzio et al., 2006; Schafer, Caetano, & Clark, 2002). Yet other investigations have found that gender does not play a significant role in the prediction of interpartner concordance for IPV or PA (Caetano et al., 2009; O'Leary & Williams, 2006). Thus, gender and victim/perpetrator status are not strong or consistent sources of reporting bias. Instead, psychological characteristics specific to the individuals under study may prove more useful.

Relationship satisfaction is one psychological characteristic that may greatly influence interpartner reporting concordance. Not surprisingly, relationship satisfaction is negatively correlated with the experience of IPV and PA (e.g., Rosen, Parmley, Knudson, & Fancher, 2002). However, victims of IPV and PA vary greatly in their experience of relationship satisfaction, with at least 27% of victims of IPV reporting above average (i.e., "excellent") levels of relationship satisfaction (O'Leary et al., 1989; Williams & Frieze, 2005). Relationship satisfaction is consistently inversely associated with attributions of partner responsibility and blame (Bradbury & Fincham, 1990) such that individuals who are satisfied with their relationship are motivated to believe good things about their partners and help their partners to look good (e.g., Fincham, Bradbury, Arias, Byrne, & Karney, 1997), whereas individuals who are not satisfied with their relationship blame their partners for negative relationship events (e.g., Byrne & Arias, 1997). Thus, high relationship satisfaction and the associated altered perception of relationship events, denial, or defensiveness may lead to the reporting of less of one's partner's IPV and PA than the partner reports.

(Langhinrichsen-Rohling & Vivian, 1994). Similarly, low relationship satisfaction and the associated attributions of partner blame may lead to the reporting of more of one's partner's IPV and PA than the partner reports.

The impact of relationship satisfaction on IPV and/or PA reporting concordance has been examined using the Conflict Tactics Scale (CTS; Straus, 1979) by Panuzio and colleagues (2006), an extended version of the CTS by Langhinrichsen-Rohling and Vivian (1994), and the Revised Conflict Tactics Scale (CTS2; Straus, Hamby, Boney-McCoy, & Sugarman, 1996) by Simpson and Christensen (2005). However, results are conflicting. Consistent with the literature reviewed above, among a sample of couples seeking marital therapy, Langhinrichsen-Rohling and Vivian (1994) found that wives' relationship satisfaction was higher among wives who reported less husband-perpetrated IPV than their husbands reported, compared to wives who reported more husband-perpetrated IPV than their husbands reported. In contrast, when examining relationship satisfaction at the level of the couple (i.e., partners' average satisfaction level), two additional studies of therapy-seeking couples did not find an association between couples' relationship satisfaction and IPV or PA reporting concordance (Panuzio et al., 2006; Simpson & Christensen, 2005).

Although these inconsistent results may be due to the use of different versions of the CTS, it is also possible that they are due to differing methods for examining the impact of relationship satisfaction on reporting concordance. That is, examining individual-level variables, rather than couple-level variables, may be a more sensitive means of identifying factors influencing interpartner reporting concordance. Unfortunately, researchers rarely examine the relative contribution of each partner's characteristics in the prediction of IPV and PA reporting concordance, despite the fact that each partner contributes equally to the couple's concordance rate. In addition, with few exceptions, researchers have not examined the direction in which the individual-level variables contribute to low within-couple reporting concordance (i.e., whether the individual-level variables are associated with reporting more or less aggression than one's partner reports). Finally, prior analytic methods do not take the level of IPV or PA into account in order to discover factors that are associated with reporting concordance beyond the impact of the level of IPV or PA. That is, because individuals with high relationship satisfaction typically experience and perpetrate lower frequencies of IPV and PA than individuals with low relationships satisfaction (Rosen et al., 2002), the impact of relationship satisfaction on reporting concordance may be an artifact of IPV/PA level such that couples with high levels of IPV/PA have more room for cross-partner differences in their reports.

To improve upon prior methodologies, we used multilevel modeling to account for the level of IPV and PA reported, as well as the dependency across partners, when examining the influence of each partner's individual characteristics on the direction of one's report of IPV and PA in comparison to one's partner's reports. We expected that interpartner IPV and PA reporting concordance rates would be in the low to moderate range. In addition, we expected that, for both husbands and wives, relationship satisfaction would be associated with low interpartner concordance such that higher relationship satisfaction would be associated with less reporting of one's partner's IPV and PA than that reported by the partner, whereas lower relationship satisfaction would be associated with more reporting of one's partner's IPV and PA than that reported by the partner.

Method

IPV and PA reporting concordance was examined using existing data from two community samples, allowing us to examine the generalizability of results across samples. Both samples included couples who were married and/or living together as if married. (Although some

couples were unmarried, participants will be referred to as “husbands” and “wives” for the sake of brevity.) The samples were similar in many ways (e.g., demographics, method and location of recruitment), but differed in terms of mean level of husband and wife IPV perpetration.

Participants

Sample 1 ($N = 88$ couples) and Sample 2 ($N = 164$ couples) consisted of heterosexual couples recruited from a midwestern metropolitan area. Husbands were in their mid-30s (Sample 1: $M = 37.1$, $SD = 9.4$ years; Sample 2: $M = 35.6$, $SD = 9.3$ years), as were wives (Sample 1: $M = 34.3$, $SD = 8.6$ years; Sample 2: $M = 33.8$, $SD = 8.9$ years). Most participants self-identified as Caucasian (Sample 1: 80% of husbands, 81% of wives; Sample 2: 75% of husbands, 79% of wives) or African American (Sample 1: 17% of husbands, 18% of wives; Sample 2: 21% of husbands, 21% of wives). Husbands' average monthly income was similar across samples (Sample 1: $M = \$2,248$, $SD = \$1,939$; Sample 2: $M = \$2,150$, $SD = \$1,161$), as was wives' monthly income (Sample 1: $M = \$1,205$, $SD = \$1,039$; Sample 2: $M = \$1,232$, $SD = \$927$). On average, husbands had some post-high school formal education (Sample 1: $M = 13.0$, $SD = 2.3$ years; Sample 2: $M = 14.1$, $SD = 2.3$ years), as did wives (Sample 1: $M = 14.1$, $SD = 2.4$ years; Sample 2: $M = 14.2$, $SD = 2.5$ years). The mean length of the couples' relationships was 9.7 ($SD = 7.1$) years in Sample 1 and 9.5 ($SD = 8.0$) years in Sample 2. On average, couples had one child together (Sample 1: $M = 1.1$, $SD = 1.5$ children; Sample 2: $M = 1.1$, $SD = 1.4$ children).

Procedures

For both samples, couples were recruited using a variety of newspaper advertisements and flyers targeted toward couples experiencing a wide range of marital satisfaction, IPV, and PA (e.g., ads asked for “happy and unhappy couples,” “couples experiencing problems and severe arguments,” “couples considering divorce”). In Sample 2, the flyers additionally specified that couples were being recruited for “a study of husbands,” and couples were oversampled for relationship distress and IPV by posting flyers in locations where they were likely to be seen by relevant populations (e.g., therapists, divorce and criminal lawyers). Approximately 10% of couples in Sample 2 were recruited from such referral sources.

When potential participants initially called the laboratory, each partner independently and privately completed a telephone screening interview including measures of demographics, relationship satisfaction, IPV, and PA. Some couples were not accepted into the study if they did not meet definitions used to form maritally distressed versus nondistressed participant groups or violent versus nonviolent participant groups used for a portion of the larger study not included in the current report. That is, couples were screened out of the study if one partner scored above 100 on the Marital Adjustment Test (MAT; Locke & Wallace, 1959) and the other partner scored below 100 but not below 80 ($n = 18$ in Sample 1, $n = 17$ in Sample 2) or if, on the CTS (Straus, 1979), either partner reported husband IPV in the past 5 years but not in the past year ($n = 18$ in Sample 1, $n = 28$ in Sample 2). Additional participants who contacted the laboratory but did not participate in the study did so because they were not interested, did not return our calls, or provided faulty contact information ($n = 74$ in Sample 1, $n = 380$ in Sample 2), or they did not meet demographic requirements (e.g., married or living together, able to read and write in English, over age 18; $n = 36$ in Sample 1, $n = 40$ in Sample 2). Those who received and accepted an invitation to complete the study attended two to three separate assessment sessions, usually scheduled 1 week apart. For the current study, participants' marital satisfaction (i.e., MAT) data from the telephone screening was used. Participants independently completed the measure of relationship aggression (i.e., CTS2) in the lab, in addition to measures and procedures not of relevance to the present study (for Sample 1, see Holtzworth-Munroe, Meehan, Herron, Rehman, &

Stuart, 2000; for Sample 2, see Clements, Holtzworth-Munroe, Schweinle, & Ickes, 2007, and Marshall & Holtzworth-Munroe, 2010). Participants were compensated at a rate of approximately \$20 per hour for their time.

Measures

Revised Conflict Tactics Scale—The CTS2 (Straus et al., 1996) is the most widely used measure of IPV, and includes physical assault, psychological aggression, sexual coercion, injury, and negotiation subscales. To be consistent with prior concordance literature and to avoid spurious results due to low base rate behaviors, only the physical assault and psychological aggression subscales were used, although the entire measure was administered. For each of 39 listed behaviors, participants indicated how many times in the past year they or their partner had engaged in the behavior. The 7-point CTS2 response scale was used including the responses of 1 (*never*), 2 (*once*), 3 (*twice*), 4 (*3–5 times*), 5 (*6–10 times*), 6 (*11–20 times*), and 7 (*more than 20 times*). The average of scores in a range was used (e.g., a response to “3–5 times” was scored as a frequency of 4 times). The physical assault subscale includes 12 items ranging from 1 (*threw something at my partner that could hurt*) to 12 (*used a knife or gun on my partner*), and the psychological aggression subscale includes 8 items ranging from 1 (*insulted or swore at my partner*) to 8 (*threatened to hit or throw something at my partner*). These subscales have good internal consistency and test–retest reliability among samples of community couples and men mandated to batterer’s intervention programs (O’Leary & Williams, 2006; Vega & O’Leary, 2007), and they have demonstrated factorial validity (Connelly, Newton, & Aarons, 2005), as well as convergent validity across a wide range of measures (e.g., Schumacher, Feldbau-Kohn, Slep, & Heyman, 2001; Schumacher, Slep, & Heyman, 2001; Straus, 2004). The CTS2 demonstrates small negative correlations with social desirability (Bell & Naugle, 2007; Straus, 2004).

Marital Adjustment Test—The MAT (Locke & Wallace, 1959) is a 15-item measure of relationship satisfaction that uses a weighted scoring system and includes questions regarding global adjustment, relationship disagreements, conflict resolution, cohesion, and communication. The MAT is one of the most frequently used measures of relationship satisfaction, and it has demonstrated split-half reliability, as well as concurrent and convergent validity. Alpha coefficients for husbands’ and wives’ reports were .78 and .80, respectively, in Sample 1, and .79 and .83, respectively, in Sample 2. Average inter-item correlations for husbands’ and wives’ reports were .26 (95% CI = .05–.44) and .26 (95% CI = .06–.45), respectively, in Sample 1, and .27 (95% CI = .13–.41) and .33 (95% CI = .19–.46), respectively, in Sample 2.

Data Analysis

Measures of Concordance—To measure reporting concordance, we used percentage of occurrence agreement, κ (Cohen, 1960), Yule’s Y (Spitznagel & Helzer, 1985), and d' prime (d'). Percentage of occurrence agreement indicates the percentage of couples who agreed that the particular form of aggression (i.e., IPV or PA) had or had not occurred. To avoid inflation of agreement due to aggression nonoccurrence, this statistic was also calculated among only those couples in which at least one partner reported the occurrence of the indicated form of aggression.

Kappa and Yule’s Y are less influenced by chance than percentage of occurrence agreement (Bartko, 1991). However, incorrect estimates may arise when base rates are skewed. In comparison to κ , the likelihood of an incorrect estimate is corrected to a degree by Yule’s Y (Hoffmann & Ninonuevo, 1994). Both of these measures are interpreted on a scale from -1 to $+1$, with 0 indicating agreement at the level of chance. Landis and Koch (1977) suggest that κ be interpreted as follows: < 0.0 = poor, $0.0–0.2$ = slight, $0.2–0.4$ = fair, $0.4–0.6$ =

moderate, 0.6–0.8 = substantial, and 0.8–1.0 = perfect. Several authors (e.g., Hoffman & Ninonuevo, 1994) interpret Yule's κ according to the established guidelines for κ .

Derived from signal detection theory, d' represents the relationship between a distribution of noise and a distribution of signal + noise. We use d' to measure the difference in couples' tendency to have only one partner indicate that a form of aggression occurred, compared to their tendency to have both partners indicate that the same form of aggression occurred. This agreement statistic is the least influenced by skewed base rates (Helberg & Brown, 1995), but it is less commonly used in the IPV/PA reporting concordance literature (for an exception, see Panuzio et al., 2006). We calculated d' by subtracting couples' standardized "false alarm" rate (i.e., the average of the probability that the husband reported aggression when the wife did not and the probability that the wife reported aggression when the husband did not) from their standardized "hit" rate (i.e., the probability of both partners reporting that aggression occurred). Higher d' scores represent both more agreement between partners (i.e., higher "hit" rates) as well as less disagreement between partners (i.e., lower "false alarm" rates). The d' statistic can be interpreted according to Cohen's (1988) guidelines for effect size d (i.e., 0.20 = small, 0.50 = medium, 0.80 = large). However, because we expect effects to be in the large range when examining reporting concordance, it is most meaningful to interpret d' statistics relative to other d' statistics (e.g., d' for concordance of IPV reports compared to d' for concordance of PA reports).

Correlates of Concordance—To most accurately measure husbands' and wives' relationship satisfaction as correlates of interpartner reporting concordance, we used multilevel modeling (HLM 6.02; Raudenbush & Bryk, 2002). To avoid inflation of agreement due to aggression nonoccurrence, all multilevel analyses were conducted including only those couples in which at least one partner reported the occurrence of the particular form of aggression of interest in the analyses (i.e., husband- or wife-perpetrated IPV or PA). Total frequencies of partners' reports of aggression were compared (i.e., comparisons were not made of dichotomized occurrence/nonoccurrence variables or item-level variables).

Because two data points per measure for each partner are needed to fit a regression line for each partner within a couple, two parallel scales were constructed for each measure of husband- and wife-perpetrated IPV and PA. This method allows us to separate measurement error from true score variance associated with the dependent variables, thus avoiding attenuation of the correlation between partners' reports and increasing the explanatory power of the models. To maintain similar variance and reliability for each parallel scale, items for each scale were matched on their standard deviations, then one item from each pair was randomly assigned to a separate scale. In Sample 1, for husband- and wife-perpetrated IPV, the coefficient α ranged from .60 to .74 and .65 to .77, respectively. For husband- and wife-perpetrated PA, the coefficient α ranged from .52 to .70 and .48 to .63, respectively. In Sample 2, for husband- and wife-perpetrated IPV, the coefficient α ranged from .79 to .87 and .59 to .79, respectively. For husband- and wife-perpetrated PA, the coefficient α ranged from .66 to .72 and .59 to .68, respectively.

The Level 1 equation consisted of the following:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{partner}) + r_{ij}$$

Y_{ij} is the report of the particular form of aggression (i.e., husband- or wife-perpetrated IPV or PA) made by partner i in couple j , with $i = 1, \dots, 4$ subscale scores per couple and $j = 1, \dots, 88$ couples in Sample 1 and $1, \dots, 164$ couples in Sample 2. Y_{1j} and Y_{2j} are the subscale scores associated with the wife in each couple and Y_{3j} and Y_{4j} are the subscale scores associated with the husband in each couple. β_{0j} is the mean of the indicated form of

aggression for couple j when $\beta_{1j} = 0$. The partner term was coded as $-.5$ for wives and $+.5$ for husbands, thus the couple average is provided when partner = 0. β_{1j} is the slope, indicating the difference in reports of aggression between partners in couple j , thus allowing us to determine the direction of disagreement (i.e., whether husbands or wives reported less of the indicated form of aggression). Note that because the relation between each parallel scale (which includes only half the CTS2 items) is used in each analysis, the means and differences between partners' reports are roughly half of the actual CTS2 subscale scores, in which frequencies for all items are summed. The residual (i.e., unpredicted) variance in Y_{ij} is represented by r_{ij} .

We also calculated Level 1 residual terms, μ_{0j} and μ_{1j} , to represent the unexplained between-couple variance in mean reports and concordance, respectively. Significant residual terms indicate significant variation across couples to be explained in the Level 2 model. We examined the association (r) between couples' mean and concordance to observe possible decreases in concordance as mean level of aggression increases. We also examined the reliabilities (ρ) of the mean and concordance coefficients. Reliability coefficients (ρ) represent the proportion of observed variance to true variance in each parameter.

In the Level 2 equation, for each couple, β_{0j} (couple mean) and β_{1j} (couple concordance) were predicted by grand mean-centered Level 2 variables (i.e., husband and wife relationship satisfaction), plus residuals. Grand mean centering was used so that the Level 2 coefficients can be interpreted relative to the mean of the outcome variables across partners (i.e., when partner = 0). The Level 2 equations consisted of the following:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} (\text{husband MAT}) + \gamma_{02} (\text{wife MAT}) + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} (\text{husband MAT}) + \gamma_{12} (\text{wife MAT}) + \mu_{1j}$$

In this example, γ_{01} and γ_{02} represent the influence of each partner's MAT score on couples' mean aggression report, while γ_{11} and γ_{12} represent the influence of these variables on the concordance between partners' aggression reports. The coefficient for each Level 2 variable, γ , is interpreted in the same manner as an unstandardized β coefficient in multiple regression (i.e., for every 1-unit increase in the explanatory variable, there is a particular unit change in couples' average ratings and in concordance within couples). Thus, the Level 2 models are similar to simultaneous multiple regression, such that the unique variance in each dependent variable is accounted for by several variables, accounting for the effect of the other variables.

Due to the different degrees of power between samples, the decreased likelihood of Type I errors using multilevel modeling, and our ability to examine hypotheses across two samples and genders, we placed substantial emphasis on effect size estimates in addition to conventional p -value estimates from null hypothesis statistical testing. We used Cohen's (1988) guidelines for interpretation of effect size d (i.e., 0.20 = small, 0.50 = medium, 0.80 = large).

Results

Descriptive Statistics

In Sample 1, at least 1 partner reported past-year husband-perpetrated IPV in each of 48 couples (54.6%), and at least 1 partner reported past-year wife-perpetrated IPV in each of 51 couples (58.0%). Based on the highest report given by either partner, these couples reported 1 to 250 acts of past-year husband-perpetrated IPV ($Mdn = 50.0$, $M = 73.9$, $SD = 70.0$), and 1 to 252 acts of past-year wife-perpetrated IPV ($Mdn = 35.0$, $M = 51.6$, $SD = 57.1$). At least 1 partner in each of 83 couples (94.3%) reported the occurrence of past-year husband-

perpetrated PA, and at least 1 partner in each of 85 couples (96.6%) reported the occurrence of past-year wife-perpetrated PA. Based on the highest report given by either partner, these couples reported 2 to 200 acts of past-year husband-perpetrated PA ($Mdn = 49.0$, $M = 50.7$, $SD = 37.8$), and 1 to 125 acts of past-year wife-perpetrated PA ($Mdn = 44.0$, $M = 49.5$, $SD = 34.2$).

In Sample 2, at least 1 partner reported past-year husband-perpetrated IPV in each of 102 couples (63.0%), and at least 1 partner reported past-year wife-perpetrated IPV in each of 94 couples (58.0%). Based on the highest report given by either partner, these couples reported 1 to 190 acts of past-year husband-perpetrated IPV ($Mdn = 6.5$, $M = 19.7$, $SD = 34.1$), and 1 to 132 acts of past-year wife-perpetrated IPV ($Mdn = 8.0$, $M = 15.3$, $SD = 20.4$). At least 1 partner in each of 162 couples (99%) reported the occurrence of past-year husband-perpetrated PA, and at least 1 partner in each of 162 couples (99%) reported the occurrence of past-year wife-perpetrated PA. Based on the highest report given by either partner, these couples reported 1 to 190 acts of past-year husband-perpetrated PA ($Mdn = 43.5$, $M = 49.9$, $SD = 40.9$), and 1 to 169 acts of past-year wife-perpetrated PA ($Mdn = 39.0$, $M = 47.1$, $SD = 36.0$).

Overall Partner Concordance

As indicated in Table 1, in Sample 1, only 57 and 45% of aggressive couples (77 and 68% of all couples) agreed regarding the occurrence of husband- or wife-perpetrated IPV, respectively. Couples' percent agreement rates were higher for husband- and wife-perpetrated PA (ranging from 83 to 87%). In Sample 2, percent agreement rates were higher than Sample 1 across all forms of aggression, ranging from 67 to 93%. Measures of κ and Yule's Y indicated that, in Sample 1, agreement was slight to moderate, and somewhat better for husband- than wife-perpetrated aggression. In Sample 2, agreement for husband- and wife-perpetrated IPV was moderate to substantial, whereas it was slight to moderate for PA. The measures of d' indicated that, in Sample 1, concordance was similar across husband- and wife-perpetrated aggression, although notably higher for PA than IPV. In Sample 2, the pattern of effects was similar as that for Sample 1, but with somewhat higher levels of agreement.

Multilevel Modeling of Concordance in Sample 1—In Sample 1, the residual terms for each Level 1 model were significant, indicating that significant variation existed across couples to be explained by the Level 2 models. In addition, all correlations between couples' mean reports of aggression (β_0) and concordance (β_1) were in the negative direction, indicating that couples' higher reports of aggression were related to husbands reporting less aggression than their wives (i.e., $r = -.39$ and $r = -.04$ for husband- and wife-perpetrated IPV, respectively; $r = -.69$ and $r = -.49$ for husband- and wife-perpetrated PA, respectively).

Husband-perpetrated IPV: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .89 and .87, respectively. As indicated in the top portion of Table 2, across couples in which at least one partner reported husband-perpetrated IPV, an average of approximately 48 acts of past-year husband-perpetrated IPV were reported (i.e., $\beta = 24.08$ multiplied by 2 due to the use of parallel scales). On average, husbands reported approximately 22 fewer such acts of IPV than wives reported (i.e., $\beta = -11.14$ multiplied by 2 due to the use of parallel scales). This effect approached statistical significance and was represented by a medium effect size ($d = .57$). Unexpectedly, neither partner's relationship satisfaction was associated with couples' mean or differential reports of husband-perpetrated IPV.

Wife-perpetrated IPV: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .71 and .70, respectively. Across couples in which at least one partner reported wife-perpetrated IPV, an average of approximately 32 ($\beta = 16.27$ multiplied by 2) acts of wife-perpetrated IPV were reported. On average, partners did not report significantly different levels of wife-perpetrated IPV ($\beta = -3.82$, $SE = 4.59$). However, husbands' relationship satisfaction was significantly associated with couples' concordance of reports, as represented by a large effect size ($d = .83$). This effect was in the expected direction such that husbands' higher relationship satisfaction was associated with husbands reporting less wife-perpetrated IPV than their wives reported, whereas husbands' lower relationship satisfaction was associated with husbands reporting more wife-perpetrated IPV than their wives reported.

Husband-perpetrated PA: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .77 and .48, respectively. Across couples in which at least one partner reported husband-perpetrated PA, an average of approximately 37 ($\beta = 18.62$ multiplied by 2) acts of husband-perpetrated PA were reported. Wives' relationship satisfaction was significantly negatively associated with couples' mean reports of husband-perpetrated PA, as represented by a medium effect size ($d = .44$). On average, husbands reported approximately 11 ($\beta = -5.66$ multiplied by 2) fewer acts of such aggression than their wives reported. Wives' relationship satisfaction was significantly associated with couples' concordance of reports, as represented by a medium effect size ($d = .57$). This effect was in the expected direction such that wives' higher relationship satisfaction was associated with wives reporting less husband-perpetrated PA than their husbands reported, whereas wives' lower relationship satisfaction was associated with wives reporting more husband-perpetrated PA than their husbands reported. In addition, the effect of husbands' relationship satisfaction on couples' concordance of reports approached statistical significance and was represented by a small to medium effect size ($d = .38$). The direction of this effect indicated that husbands' higher relationship satisfaction was associated with husbands reporting less husband-perpetrated PA than their wives reported, whereas husbands' lower relationship satisfaction was associated with husbands reporting more husband-perpetrated PA than their wives reported.

Wife-perpetrated PA: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .61 and .32, respectively. Across couples in which at least one partner reported wife-perpetrated PA, an average of approximately 48 ($\beta = 24.20$ multiplied by 2) acts of wife-perpetrated PA were reported. Wives' relationship satisfaction was significantly negatively associated with couples' mean reports of wife-perpetrated PA, as represented by a large effect size ($d = .86$). On average, partners did not report significantly different levels of wife-perpetrated PA ($\beta = -2.74$, $SE = 2.74$). However, husbands' relationship satisfaction was significantly associated with couples' concordance of reports, as represented by a large effect size ($d = 1.08$). This effect was in the expected direction such that husbands' higher relationship satisfaction was associated with husbands reporting less wife-perpetrated PA than their wives reported, whereas husbands' lower relationship satisfaction was associated with husbands reporting more wife-perpetrated PA than their wives reported. In addition, wives' relationship satisfaction was significantly associated with couples' concordance of reports, as represented by a medium to large effect size ($d = .62$). The direction of this effect indicated that wives' higher relationship satisfaction was associated with wives reporting less wife-perpetrated PA than their husbands reported, whereas wives' lower relationship satisfaction was associated with wives reporting more wife-perpetrated PA than their husbands reported.

Multilevel Modeling of Concordance in Sample 2—In Sample 2, the residual terms for each Level 1 model were significant, indicating that significant variation existed across couples to be explained by the Level 2 models. Correlations between couples' mean reports of husband aggression (β_0) and concordance (β_1) were of low magnitude (i.e., $r = -.01$ for IPV; $r = .05$ for PA). Correlations between couples' mean reports of wife aggression and concordance were in the positive direction (i.e., $r = .57$ for IPV; $r = .18$ for PA), suggesting that higher couples' mean reports were associated with husbands reporting more such aggression than their wives.

Husband-perpetrated IPV: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .96 and .78, respectively. As indicated in the lower portion of Table 2, across couples in which at least one partner reported husband-perpetrated IPV, an average of approximately 14 ($\beta = 7.05$ multiplied by 2) acts of past-year husband-perpetrated IPV were reported. On average, partners did not report significantly different levels of husband-perpetrated IPV ($\beta = -1.41$, $SE = 1.05$). However, wives' relationship satisfaction was significantly associated with couples' concordance of reports, as represented by a medium effect size ($d = .48$). This effect was in the expected direction such that wives' higher relationship satisfaction was associated with wives reporting less husband-perpetrated IPV than their husbands reported, whereas wives' lower relationship satisfaction was associated with wives reporting more husband-perpetrated IPV than their husbands reported.

Wife-perpetrated IPV: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .92 and .70, respectively. Across couples, an average of approximately 11 ($\beta = 5.57$ multiplied by 2) acts of wife-perpetrated IPV were reported. Wives' relationship satisfaction was significantly negatively associated with couples' mean reports of wife-perpetrated IPV, as represented by a medium effect size ($d = .46$). On average, wives reported approximately three ($\beta = 1.40$ multiplied by 2) fewer such acts of wife-perpetrated IPV than husbands reported. This effect approached statistical significance and was represented by a small to medium effect size ($d = .38$). Unexpectedly, neither partner's relationship satisfaction was associated with couples' concordance of reports of wife-perpetrated IPV.

Husband-perpetrated PA: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .85 and .60, respectively. Across couples in which at least one partner reported husband-perpetrated PA, an average of approximately 38 ($\beta = 19.04$ multiplied by 2) acts of husband-perpetrated PA were reported. Husbands' and wives' relationship satisfaction were both significantly negatively associated with couples' mean reports of husband-perpetrated PA, as represented by a medium ($d = .50$) and medium to large effect size ($d = .63$), respectively. On average, husbands reported approximately six ($\beta = -2.90$ multiplied by 2) fewer acts of such aggression than their wives reported. Wives' higher relationship satisfaction was significantly associated with couples' concordance of reports, as represented by a medium effect size ($d = .57$). This effect was in the expected direction such that wives' higher relationship satisfaction was associated with wives reporting less husband-perpetrated PA than their husbands reported, whereas wives' lower relationship satisfaction was associated with wives reporting more husband-perpetrated PA than their husbands reported. In addition, the effect of husbands' relationship satisfaction on couple's concordance of reports approached statistical significance and was represented by a small to medium effect size ($d = .31$). The direction of this effect indicated that husbands' higher relationship satisfaction was associated with husbands reporting less husband-perpetrated PA than their wives reported, whereas husbands' lower relationship satisfaction

was associated with husbands reporting more husband-perpetrated PA than their wives reported.

Wife-perpetrated PA: Reliability estimates for couples' mean (β_0) and concordance (β_1) of reports in the Level 2 model were .83 and .58, respectively. Across couples in which at least one partner reported wife-perpetrated PA, an average of approximately 36 ($\beta = 17.88$ multiplied by 2) acts of wife-perpetrated PA were reported. Husbands' and wives' relationship satisfaction were both significantly negatively associated with couples' mean reports of wife-perpetrated PA, as represented by medium effect sizes ($d = .46$ and $d = .47$, respectively). On average, partners did not report significantly different levels of wife-perpetrated PA ($\beta = -0.73$, $SE = 1.25$). However, husbands' relationship satisfaction was significantly associated with couples' concordance of reports, as represented by a small to medium effect size ($d = .34$). As expected, husbands' higher relationship satisfaction was associated with husbands reporting less wife-perpetrated PA than their wives reported, whereas husbands' lower relationship satisfaction was associated with husbands reporting more wife-perpetrated PA than their wives reported. In addition, the effect of wives' relationship satisfaction on couples' concordance of reports approached statistical significance and was represented by a small effect size ($d = .26$). The direction of this effect indicated that wives' higher relationship satisfaction was associated with wives reporting less wife-perpetrated PA than their husbands reported, whereas wives' lower relationship satisfaction was associated with wives reporting more wife-perpetrated PA than their husbands reported.¹

Discussion

Similar to other areas of study in which intimate partners report on the same behaviors (e.g., Jacobson & Moore, 1981; Mikelson, 2008), low to moderate levels of concordance between partners in the reporting of IPV and PA have been widely documented (Archer, 1999; Caetano et al., 2009). The current study provides another example of inadequate levels of interpartner concordance of IPV and PA reports among two community samples, extending the substantial body of research on interpartner reporting concordance on the CTS (Straus, 1979) among community samples to the CTS2 (Straus et al., 1996). The current results also support the findings of prior studies of interpartner reporting concordance on the CTS2 among clinical samples (O'Leary & Williams, 2006; Simpson & Christensen, 2005) and a sample of construction industry union members (Cunradi, Bersamin, & Ames, 2009).

In line with prior research (Archer, 1999; Caetano et al., 2009; O'Leary & Williams, 2006; Panuzio et al., 2006; Perry & Fromuth, 2005; Schafer et al., 2002), we also examined the suggestion that the direction of IPV and PA reporting, compared to one's partner, depends upon the gender and/or victim/perpetrator status of the reporter. Consistent with this body of research, gender and victim/perpetrator status did not regularly predict partners' relative reporting of IPV and PA. However, there was some suggestion that victim/perpetrator status may be a stronger predictor of reporting patterns as the frequency of aggression increases. That is, in Sample 1, husbands reported less of their own IPV and PA perpetration than their wives reported and this difference between partners became larger as the frequency of couples' mean reports of aggression increased. In addition, in Sample 2, wives reported somewhat less of their own IPV perpetration than their husbands reported and this difference

¹We also examined all multilevel models when including all couples, rather than restricting the analyses to those couples who reported the indicated form of aggression. In Sample 1, the pattern of results remained the same. In Sample 2, husbands' relationship satisfaction became significantly associated with couples' concordance for reports of husband-perpetrated IPV, $t(161) = -2.28$, $p < .05$, $d = .36$. The direction of this effect indicated that husbands' higher relationship satisfaction was associated with husbands reporting less husband-perpetrated IPV than their wives reported, whereas husbands' lower relationship satisfaction was associated with husbands reporting more husband-perpetrated IPV than their wives reported.

between partners became larger as the frequency of couples' mean reports of aggression increased. Regardless, the pattern of results was not reliable. Instead, examination of individual psychological variables that characterize the population of interest may be a more fruitful method of determining factors that impact concordance of reports.

Compared to gender and victim/perpetrator status, individual-level relationship satisfaction provided the clearest picture of an important factor that influences reporting of IPV and PA. As expected, for both husbands and wives, higher relationship satisfaction was associated with reporting less of one's partner's IPV and PA than that reported by the partner, whereas lower relationship satisfaction was associated with reporting more of one's partner's IPV and PA than that reported by the partner. This pattern of results was consistently found across samples for the perpetration of PA, but meaningful effects were found in only one sample for the perpetration of IPV (i.e., Sample 1 for the prediction of wife-perpetrated IPV and Sample 2 for the prediction of husband-perpetrated IPV). The greater consistency and robustness of the effect for PA compared to IPV may be due to PA behaviors (e.g., yelling, saying something to spite one's partner) being more open to interpretation than the more behaviorally anchored IPV behaviors (e.g., slap, burn, kick), thus allowing schema-based subjectivity to impact the reporting of such behaviors.

The general pattern of these results supports the work of Langhinrichsen-Rohling and Vivian (1994), in the only other study to examine relationship satisfaction as an individual-level variable. Given that relationship satisfaction is inversely associated with attributions of partner responsibility and blame (Bradbury & Fincham, 1990), such attributions may lead individuals who are not satisfied with their relationships to report more of their partners' IPV and PA than their partners report. Similarly, individuals who are satisfied with their relationships may be motivated to believe good things about their partners and help their partners to look good (Fincham et al., 1997), leading such individuals to be reluctant to report their partners' aggression in the context of an otherwise happy relationship. However, throughout this report, many of our interpretations are contingent upon the assumption that the higher report of aggression is the more accurate report, an assumption that simply cannot be addressed in this form of research. Alternative interpretations are warranted. For example, it is possible that individuals who are not satisfied with their relationship may provide inaccurately exaggerated estimates of the frequency of their partner's aggression. In any case, given the cognitive tasks inherent in the completion of measures such as the CTS2 (e.g., summing the number of behaviors that occur over a relatively long period of time), it may be that general cognitive biases (e.g., a schema of a happy or distressed relationship) play a major role in the estimation of behavior frequencies.

Although not predicted, the current data also suggest that higher relationship satisfaction may be associated with reporting less of one's own PA than that reported by one's partner, whereas lower relationship satisfaction may be associated with reporting more of one's own PA than that reported by one's partner. These results typically did not reach statistical significance and did not occur for the prediction of IPV concordance; thus they should be interpreted extremely cautiously. However, these results are consistent with the possibility that one's general schema regarding their relationship being satisfying or dissatisfying may also impact partners' reporting of their own behavior. Thus, although overlapping considerably, satisfaction with one's relationship may play a greater role in the reporting of IPV and PA than satisfaction with one's partner. This may be a function of how individuals remember relationship conflicts and how the partnership functions to resolve conflicts. Together, these results suggest that the negative correlations that have been consistently found between relationship satisfaction and IPV and PA (Rosen et al., 2002; Schumacher, Feldbau-Kohn, et al., 2001; Schumacher, Slep, et al., 2001) may be partially inflated by partners' satisfaction-based reporting bias.

In terms of the assessment of relationship behaviors more generally, the results of the current study imply that the impact of individual-level variables on interpartner reporting concordance may be erased when aggregating across partners in a couple. That is, it is possible that Panuzio et al. (2006) and Simpson and Christensen (2005) may have found that relationship satisfaction was associated with concordance of reports if they had not examined relationship satisfaction at the level of the couple. As researchers begin to examine additional factors that impact couples' low IPV and PA reporting concordance, unless deliberately based on a theoretical rationale, it will be important to examine such factors at the individual level and to examine the direction in which such factors impact concordance. Further, studies that have examined other potential predictors of IPV and PA reporting concordance have not included assessment of relevant constructs among both partners (e.g., researchers have examined only one partner's alcohol problems as a predictor of IPV and PA reporting concordance; Panuzio et al., 2006). This is surprising given correlations between partners on many variables used to predict concordance (e.g., Leonard & Das Eiden, 1999). The current study methods may be used as an improved base for examining other predictors of concordance across each partner in a couple.

The results of the current study should be considered in light of its limitations. Because some couples were screened out of participation if they did not meet participant group criteria used in the larger studies (i.e., the husband violent vs. nonviolent and maritally distressed vs. nondistressed criteria described in the Procedures section), the pooling of these groups may have produced upwardly biased regression estimates. In addition, because we screened some couples out of participation based on husband-perpetrated IPV but not wife-perpetrated IPV, our estimates of concordance for husband and wife IPV may not be fully comparable. However, given the small number of couples screened out of participation for these reasons (particularly compared to other reasons for nonparticipation) and our limiting the primary analyses to couples in which at least one person reported the indicated form of aggression, we expect the impact of this limitation to be minimal.

Additional caveats bear note. First, findings from this and all investigations of IPV and PA reporting concordance are assuaged by the fact that the true level of aggression in a relationship is not known. Future investigations should attempt to address the assumption that higher reports are more accurate, which underlies much of the existing IPV and PA concordance research. Second, because recruitment of both samples focused on oversampling community couples experiencing relationship difficulties, results may not be fully generalizable to treatment-seeking or court-referred populations, or the community at large. Notably, rates of IPV and PA in the current study generally were more consistent with those documented among couples seeking marital therapy (e.g., O'Farrell, Murphy, Stephan, Fals-Stewart, & Murphy, 2004) than samples of couples drawn from the community (e.g., Panuzio & DiLillo, in press) in which considerably lower rates of aggressive behavior, particularly IPV, are typically reported. Although Langhinrichsen-Rohling and Vivian (1994) obtained similar results among a clinic sample, other factors such as a desire to express the severity of the aggression to a therapist or the influence of other clinical conditions (e.g., depression, antisocial personality characteristics) may more strongly impact reporting behavior among clinical samples. Further, the strength or pervasiveness of the observed phenomena among unselected community samples, particularly those with lower rates of IPV and PA, is not known. Finally, the results of the current study are specifically relevant to the CTS2. Although our use of a more psychometrically sound measure (compared to the vast majority of studies of interpartner reporting concordance that have used the original CTS) should have increased our ability to disentangle a true correlate of interpartner reporting concordance from measurement error associated with the original CTS, the generalizability of these findings to other measures of IPV and PA is not known. Alternative IPV assessment methods that differ in reporting context (e.g., the Timeline

Followback Spousal Violence interview; Fals-Stewart, Birchler, & Kelley, 2003) may possess a distinctly different set of correlates of concordance.

Despite these caveats, the present study provides an example of an advanced statistical methodology through which variables that impact poor interreporter reliability can be investigated. We examined individuals' relationship satisfaction as one important variable that may impact IPV and PA reporting biases. However, this predictor is multifaceted and future research may benefit from an examination of basic processes or mechanisms more directly associated with the reporting of behaviors over a relatively long period of time under intense emotional arousal (e.g., relationship schemas and attributions). Research on IPV and PA has increased substantially in recent years (Langhinrichsen-Rohling, 2005), and will not progress in an ideal fashion without a better understanding of the reliability of the primary measurement device. Future investigations should continue to investigate psychological, dyadic, and methodological correlates of IPV and PA reporting concordance, as well as the concordance of reports for other seemingly objective behaviors, in order to shed light on this important assessment issue. With a solid knowledge base of the factors that impact reporting concordance, researchers and clinicians can enhance reliability and validity of assessment instruments by countering these systematic biases. Until such work is completed, we do not advise clinicians working with couples to interpret the current findings as suggesting that distressed partners report more aggression than actually occurs. Remembering that the true level of aggression is not known, all reports should be taken seriously and responded to as valid, and potentially even underestimated, representations of each person's subjective reality.

Acknowledgments

Dr. Marshall is supported by the National Institutes of Health's Building Interdisciplinary Research Careers in Women's Health (BIRCWH) program (1 K12 HD055882).

References

- Archer J. Assessment of the reliability of the conflict tactics scales: A meta-analytic review. *Journal of Interpersonal Violence*. 1999; 14:1263–1289.
- Armstrong TG, Wernke JY, Medina KL, Schafer J. Do partners agree about the occurrence of violence? A review of the current literature. *Trauma, Violence, and Abuse*. 2002; 3:181–193.
- Bartko JJ. Measurement and reliability: Statistical thinking considerations. *Schizophrenia Bulletin*. 1991; 17:483–489. [PubMed: 1947873]
- Bell KM, Naugle AE. Effects of social desirability on students' self-reporting of partner abuse perpetration and victimization. *Violence and Victims*. 2007; 22:243–256. [PubMed: 17479559]
- Bradbury TN, Fincham FD. Attributions in marriage: Review and critique. *Psychological Bulletin*. 1990; 107:3–33. [PubMed: 2404292]
- Byrne CA, Arias I. Marital satisfaction and marital violence: Moderating effects of attributional processes. *Journal of Family Psychology*. 1997; 11:188–195.
- Caetano R, Field C, Ramisetty-Mikler S, Lipsky S. Agreement on reporting of physical, psychological, and sexual violence among white, black, and Hispanic couples in the United States. *Journal of Interpersonal Violence*. 2009; 24:1318–1337. [PubMed: 18768744]
- Clements K, Holtzworth-Munroe A, Schweinle W, Ickes W. Empathic accuracy of intimate partners in violent versus nonviolent relationships. *Personal Relationships*. 2007; 14:369–388.
- Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 1960; 10:37–46.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum; 1988.

- Connelly CD, Newton RR, Arons GA. A psychometric examination of English and Spanish versions of the revised conflict tactics scales. *Journal of Interpersonal Violence*. 2005; 20:1560–1579. [PubMed: 16246917]
- Cunradi CB, Bersamin M, Ames G. Agreement on intimate partner violence among a sample of blue-collar couples. *Journal of Interpersonal Violence*. 2009; 24:551–568. [PubMed: 18430971]
- Fals-Stewart W, Birchler GR, Kelley ML. The timeline followback spousal violence interview to assess physical aggression between intimate partners: Reliability and validity. *Journal of Family Violence*. 2003; 18:131–142.
- Fincham FD, Bradbury TN, Arias I, Byrne CA, Karney BR. Marital violence, marital distress, and attributions. *Journal of Family Psychology*. 1997; 11:367–372.
- Helberg, CP.; Brown, RL. Concordance in fourfold tables: A simulation study of statistics indices of agreement. Paper presented at the meeting of the International Applied Statistics in Industry Conference; Dallas, TX. 1995 Jun.
- Hoffmann NG, Ninonuevo FG. Concurrent validation of substance abusers self-reports against collateral information: Percentage agreement versus kappa versus Yule's Y. *Alcoholism: Clinical and Experimental Research*. 1994; 18:231–237.
- Holtzworth-Munroe A, Meehan JC, Herron K, Rehman U, Stuart GL. Testing the Holtzworth-Munroe and Stuart (1994) batterer typology. *Journal of Consulting and Clinical Psychology*. 2000; 68:1000–1019. [PubMed: 11142534]
- Jacobson N, Moore D. Spouses as observers of the events in their relationships. *Journal of Consulting and Clinical Psychology*. 1981; 49:269–277. [PubMed: 7217493]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
- Langhinrichsen-Rohling J. Top 10 greatest "hits": Important findings and future directions for intimate partner violence research [Special issue]. *Journal of Interpersonal Violence*. 2005; 20:108–118. [PubMed: 15618567]
- Langhinrichsen-Rohling J, Vivian D. The correlates of spouses' incongruent reports of marital aggression. *Journal of Family Violence*. 1994; 9:265–283.
- Leonard KE, Das Eiden R. Husband's and wife's drinking: Unilateral or bilateral influences among newlyweds in a general population sample. *Journal of Studies on Alcohol*. 1999; 13:130–113.
- Locke HJ, Wallace KM. Short marital-adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living*. 1959; 21:251–255.
- Marshall AD, Holtzworth-Munroe A. Recognition of wives emotional expressions: A mechanism in the relationship between psychopathology and intimate partner violence perpetration. *Journal of Family Psychology*. 2010; 24:21–30. [PubMed: 20175605]
- Mikelson KS. He said, she said: Comparing mother and father reports of father involvement. *Journal of Marriage and Family*. 2008; 70:613–624.
- O'Farrell TJ, Murphy CM, Stephan SH, Fals-Stewart W, Murphy M. Partner violence before and after couples-based alcoholism treatment for male alcoholic patients: The role of treatment involvement and abstinence. *Journal of Consulting and Clinical Psychology*. 2004; 72:202–217. [PubMed: 15065955]
- O'Leary KD, Barling J, Arias I, Rosenbaum A, Malone J, Tyree A. Prevalence and stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology*. 1989; 57:263–268. [PubMed: 2785126]
- O'Leary KD, Williams MC. Agreement about acts of aggression in marriage. *Journal of Family Psychology*. 2006; 20:656–662. [PubMed: 17176201]
- Panuzio J, DiLillo D. Physical, psychological, and sexual intimate partner aggression among newlywed couples: Longitudinal prediction of marital satisfaction. *Journal of Family Violence*. (in press).
- Panuzio J, O'Farrell TJ, Marshall AD, Murphy CM, Murphy M, Taft CT. Intimate partner aggression reporting concordance and correlates of agreement among men with alcohol use disorders and their female partners. *Assessment*. 2006; 13:266–279. [PubMed: 16880279]
- Perry AR, Fromuth ME. Courtship violence using couple data: Characteristics and perceptions. *Journal of Interpersonal Violence*. 2005; 20:1078–1095. [PubMed: 16051728]

- Raudenbush, SW.; Bryk, AS. Hierarchical linear models: Applications and data analysis methods. 2nd ed.. Thousand Oaks, CA: Sage; 2002.
- Rosen LN, Parmley AM, Knudson KH, Fancher P. Differences in the experience of intimate partner violence among active duty and U.S. Army soldiers. *Military Medicine*. 2002; 167:959–963. [PubMed: 12502167]
- Schafer J, Caetano R, Clark C. Agreement about violence in U.S. couples. *Journal of Interpersonal Violence*. 2002; 17:457–470.
- Schumacher JA, Feldbau-Kohn S, Slep AMS, Heyman RE. Risk factors for male-to-female partner physical abuse. *Aggression and Violent Behavior*. 2001; 6:281–352.
- Schumacher JA, Slep AMS, Heyman RE. Risk factors for male-to-female partner psychological abuse. *Aggression and Violent Behavior*. 2001; 6:255–268.
- Simpson LE, Christensen A. Spousal agreement regarding relationship aggression on the conflict tactics scale–2. *Psychological Assessment*. 2005; 17:423–432. [PubMed: 16393009]
- Spitznagel EL, Helzer JE. A proposed solution to the base rate problem in kappa statistics. *Archives of General Psychiatry*. 1985; 42:725–728. [PubMed: 4015315]
- Straus MA. Measuring intrafamily conflict and violence: The conflict tactics scales. *Journal of Marriage and the Family*. 1979; 41:75–88.
- Straus MA. Cross-cultural reliability and validity of the revised conflict tactics scale: A study of university student dating couples in 17 nations. *Cross-Cultural Research*. 2004; 38:407–432.
- Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The revised conflict tactics scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*. 1996; 17:283–316.
- Vega EM, O’Leary KD. Test–retest reliability of the revised conflict tactics scales (CTS2). *Journal of Family Violence*. 2007; 22:703–708.
- Williams SL, Frieze IH. Patterns of violent relationships, psychological distress, and marital satisfaction in a national sample of men and women. *Sex Roles*. 2005; 52:771–784.

Table 1
Interpartner Concordance on the Occurrence/Nonoccurrence of IPV and PA Perpetration

Perpetrator	Aggression	Aggressive Couples		All Couples		κ	Yule's Y	d'
		% Agreement		% Agreement				
Sample 1								
Husband	IPV	56.8		76.5		.52	.53	0.69
	PA	85.9		86.6		.35	.51	2.40
Wife	IPV	44.7		67.9		.34	.35	0.76
	PA	83.1		83.5		.15	.28	2.27
Sample 2								
Husband	IPV	66.7		79.3		.59	.59	1.05
	PA	93.2		93.3		.23	.53	3.24
Wife	IPV	72.3		84.1		.68	.68	1.20
	PA	92.5		92.5		.10	.34	3.18

Note. IPV = intimate partner violence (physical assault); PA = psychological aggression.

Table 2

Level 2 Models

Fixed Effects	Intimate Partner Violence				Psychological Aggression											
	Husband-Perpetrated		Wife-Perpetrated		Husband-Perpetrated		Wife-Perpetrated									
	B	SE	t	d	B	SE	t	d	B	SE	t	d	B	SE	t	d
Sample 1																
Couple Mean																
Intercept	24.08	3.25	7.42***	2.24	16.27	2.36	6.91***	2.02	18.62	1.55	11.98***	2.70	24.20	1.78	13.57***	3.96
Husband satisfaction	0.03	0.12	0.28	.08	0.08	0.09	0.96	.28	-0.02	0.06	-0.32	.07	0.03	0.06	0.40	.12
Wife satisfaction	-0.17	0.12	-1.34	.40	-0.10	0.09	-1.51	.44	-0.12	0.06	-1.96*	.44	-0.21	0.07	-2.94**	.86
Couple Concordance																
Intercept	-11.14	5.94	-1.88†	.57	-3.82	4.59	-0.83	.24	-5.66	2.09	-2.71**	.61	-2.74	2.74	-1.00	.29
Husband satisfaction	-0.18	0.21	-0.86	.26	-0.47	0.17	-2.84**	.83	-0.14	0.08	-1.67†	.38	-0.37	0.10	-3.70***	1.08
Wife satisfaction	0.08	0.23	0.34	.10	0.15	0.18	0.96	.28	0.20	0.08	2.53*	.57	0.23	0.11	2.12*	.62
Sample 2																
Couple Mean																
Intercept	7.05	1.29	5.46***	1.10	5.57	0.76	7.35***	1.54	19.04	1.06	18.02***	2.86	17.88	0.97	18.52***	2.93
Husband satisfaction	-0.06	0.05	-1.23	.25	0.01	0.03	0.39	.08	-0.15	0.05	-3.18**	.50	-0.12	0.04	-2.91**	.46
Wife satisfaction	-0.07	0.05	-1.39	.28	-0.07	0.03	-2.18*	.46	-0.16	0.04	-4.00***	.63	-0.11	0.04	-3.00**	.47
Couple Concordance																
Intercept	-1.41	1.05	-1.35	.27	1.40	0.77	1.83†	.38	-2.90	1.28	-2.26*	.36	-0.73	1.25	-0.59	.09
Husband satisfaction	-0.06	0.04	-1.31	.26	-0.01	0.03	-0.44	.09	-0.11	0.06	-1.93†	.31	-0.12	0.05	-2.12*	.34
Wife satisfaction	0.10	0.04	2.40*	.48	-0.02	0.03	-0.70	.15	0.18	0.05	3.58***	.57	0.08	0.05	1.65†	.26

Note.

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