Assessing Conflict Tactics Scale Validity by Examining Intimate Partner Violence
Over-reporting

Jeffrey Ackerman Ph.D.

Griffith University

G06_3.07 Gold Coast Campus

Southport, Queensland 4222

Australia

Office: +61(07) 5552 8013

US voicemail: 979-446-0055

FAX: +61(07) 5552 9918

j.ackerman@griffith.edu.au

July 14, 2015

Ackerman, J. M. (2017). Assessing Conflict Tactics Scale Validity by Examining Intimate Partner Violence Over-reporting. *Psychology of Violence*, 8(2), 207-217. doi:10.1037/vio0000112

Assessing Conflict Tactics Scale Validity by Examining Intimate Partner Violence
Over-reporting

Objective: Although many scholars have questioned on a logical basis the validity of the Conflict Tactics Scale (CTS) to adequately measure intimate partner violence (IPV), relatively few critiques have used extensive empirical data gathered specifically for this purpose. This research analyzed these types of data to investigate an important mechanism of potential validity problems, IPV misreporting, by adding context questions to determine whether participants endorsed (as if actual IPV) accidents or other acts that neither party took seriously. The objective was to determine not only the extent to which this form of over-reporting occurs, but also how males and females differed in misreporting patterns. **Method:** 1,758 students from one US and one Australian university completed a computer-administered survey. Multi-level logistic regression subsequently assessed the degree to which several factors predicted whether participants over-reported CTS items. **Results:** Of the 1,174 event endorsements, 22.1% were classified as over-reports. Whether males or females were more prone to over-reporting, however, differed across event type, sample, age, relationship status, perpetration versus victimization, and current versus former partnerships. There were statistically significant interactions between gender and many of these factors. Among the most important of the findings was that males were more likely to over-report victimizations by female partners while females were more likely to over-report perpetrations against male partners. **Conclusions:** The magnitude and intricate gendered nature of the over-reporting problem imply that over-reporting is a substantial problem having the potential to negatively affect scale validity and thus the testing of partner-violence theories.

Keywords: Intimate Partner Violence, Domestic Violence, Conflict Tactics Scale, Over-reporting, Misreporting

For over four decades, many scholars have considered Straus and colleagues' (1979) conflicts tactics scale (CTS) the gold-standard for assessing intimate partner violence (Langhinrichsen-Rohling, 2005). Despite widespread use, the scale's validity continues to be questioned because several CTS-based findings contrast with those based upon alternative data sources (Hamby, 2009). A consistent critique that may explain why CTS findings differ from alternative sources has been the scale's failure to account for the context in which purported aggression occurs (DeKeseredy, 1995). Critics suggest that this failure may potentially cause the scale to mix events such as self-defense (Saunders, 1986), innocent horseplay (Lehrner & Allen, 2014), or other benign and/or self-protective acts with consequential victimization (Kimmel, 2002). These suggestions have resulted in calls for additional research using a variety of methodologies to further assess CTS validity (Follingstad & Bush, 2014). The current study has answered this call by adding to a traditional CTS scale follow-up questions to assess event context. The most basic purpose in doing so is to expand upon the author's previously published assessments about measuring event context to determine whether IPV events are over-reported by some survey participants, that is, to assess the occurrence of a specific form of misreporting involving false-positive item endorsements outside the conceptual understanding of partner violence as intended by scale design. Although a simple estimate of the degree to which IPV is over-reported is a key part of this research, the main purpose of the present study is to apply multi-level regression techniques with CTS items as the primary analytic unit to participant samples from two countries (United States and Australia) to assess which factors predict the over-reporting of various CTS items. Additional purposes are to assess whether over-reporting occurs in a gendered fashion or whether there are other systematic over-reporting errors.

A more detailed rationale for the purposes underlying this study includes the realization that while certain forms of misreporting error have the potential to over- or under-estimate IPV

prevalence, a quite problematic scenario is where misreporting occurs at different rates or in different ways for different groups of people. For example, CTS critics have suggested that men may over-report female-perpetrated aggression, while women may more accurately report the aggressive behavior of male partners (Kimmel, 2002). If so, this may explain certain controversial findings found in survey-based research, for example, according to many CTS-based studies, women and men are equally aggressive against intimate partners (Johnson, 1995) (Ackerman, 2015; Ackerman & Field, 2011).

One reason why misreporting may occur at high rates when measuring discrete acts of ostensibly violent behavior is the variety of meanings attached by various sub-populations to the terminology used by the CTS (Hamby, 2009). For example, in recent years it has become increasingly clear that emerging or young adults use terms such as “hitting” and “kicking” in ways much different to the ways older individuals use these terms. Younger individuals have often been found to use these terms to describe acts of horseplay or “low-risk sexual behaviors,” which have consequence and meaning best characterized as involving benign or even mutually pleasurable aspects (Grych & Hamby, 2014; Hamby, 2005; Lehrner & Allen, 2014).

Although the purely quantitative IPV literature has generally not addressed the degree to which research participants vary in their interpretations of CTS questions, mixed-methods approaches have used increasingly sophisticated techniques to illustrate the extent to which potentially serious over-reporting occurs. Almost thirty years ago, for example, Margolin (1987) described a case where a survey respondent endorsed an item about being “kicked” by a partner but explained during a follow-up interview that the kick occurred in a light-hearted manner that was not viewed as aggressive by either party.

Years later, in a 2014 issue of this journal, Lehrner and Allen (2014) compared quantitative CTS reports with a relatively large number of systematic follow-up interviews for

the purposes of illustrating the extent to which initial CTS item endorsements might best be characterized as playful wrestling, roughhousing, and mock violence in ways inconsistent with conceptual understandings of IPV. These scholars found that CTS scores frequently conflate play and mock aggression with meaningful violence, and when meaningful violence does occur, the scale may not adequately distinguish victims from perpetrators or mutual combatants. For example, among respondents identified as violent by the CTS, 19.2% were reporting on only playful behavior and 3.9% were reporting self-defense.

The findings of Lehrner and Allen illustrate an important, but often overlooked, distinction in the IPV literature. Their 19.2% figure did not simply represent *minor* violence reported by survey participants that *should have been reported* but better differentiated from *serious* violence by the scale, but rather *enjoyable, playful acts involving no intention to commit harm*. The former, of course, simply represents low sensitivity by the CTS to differentiate minor from serious violence, while the latter represents a potentially serious validity problem.

Lehrner and Allen provided examples that clearly illustrate the importance of making this distinction. In one instance, a research participant was classified by the CTS as a victim of severe and frequent violence due to the reporting of throwing, pushing, shoving, grabbing, and slapping, which involved two instances of minor injury. Follow-up interviews, however, revealed that the injuries were minor accidents that happened during playful behaviors that occurred when the couple was getting along quite well. In other words, they were the unequivocal misreporting of playful acts *as if they were IPV*.

An important aspect to this (and similar) examples that scholars should keep in mind is that all survey research will contain some inevitable amount of error involving incorrectly reported events. The most important question, however, is not whether we can find examples of misreports, but whether the misreports occur to a different degree among different types of

survey respondents in ways that matter for analytical purposes. In particular, what is critical to know for the testing of many theories of IPV's etiology is whether we might find that this form of misreporting occurs to a different degree among men and women.

The author's recent line of research compliments the work of Lehrner and Allen. A 2015 study relied upon both male and female participants with a greater age range than used by Lehrner and Allen to focus upon whether over-reporting varies by age and gender. In addition to confirming the findings of Lehrner and Allen by uncovering substantial over-reporting, the study also noted that males over-reported more than females. Moreover, it found that over-reporting decreased across age with nuances about these associations depending on whether perpetration or victimization events were examined.

The current study adds to both the work of Lehrner and Allen and the author's research in several ways. Most basically, the present study provides further information about how over-reporting varies by age and gender across different samples from two countries. The different samples help determine whether over-reporting patterns differ across the cultures and across different samples within the same country. In addition, the present study determines whether over-reporting rates vary across different types of purported IPV events and whether these differences occur in a gendered fashion. Most importantly, the current study focuses upon whether gender differences in over-reporting may occur in a relatively straightforward fashion that might enable scholars to apply simple corrections. For example, if males (or females) consistently over-report IPV at a known rate relative to the other gender, data analysts might be able to correct for this systematic over-reporting tendency. If, however, the male versus female over-reporting rates do not remain consistent across various samples or for various forms of IPV, simple corrective measures may be difficult if not impossible to apply.

In sum, the current study addresses the following specific research purposes: (1) a determination of whether male and female over-reporting differs across samples drawn from two countries (Australian and the United States) or across samples within the same country; (2) a determination of whether there are consistent gendered differences in the over-reporting rates for different IPV events purportedly measured by the CTS; and (3) a determination of whether other common demographic characteristics like participant age affect IPV over-reporting.

Method

Participants

Participants were students enrolled at one of two universities in either Texas in the United States or in Queensland, Australia. They were drawn using different sampling methods dictated largely by the constraints of each University's human subjects policies and the resources available to the author and his collaborators. The different samples were drawn for the purposes of testing whether they would produce different results. Potential participants were included in the data analysis if they had at least one applicable relationship after high school graduation. Same-sex partnerships, however, were excluded because there were too few for suitable results.

Usable information about 910 opposite-sex partnerships was obtained from the US University's psychology pool. Three different comparison samples were drawn at a large Australian university from traditional "in-person" as well as from "on-line" classes. The three Australian samples were drawn in different ways. The first Australian sample was obtained by sending email invitations to the University's criminal justice students who were taking required classes. The second sample from this university differed in that the email invitations were sent to students in selected elective courses. These electives contained a much higher proportion of students from alternative fields, the primary one being psychology. Together, 1,801 invitation emails were sent to students in these two groups. Participants in both samples as well as a third

from this university (described below) were offered an incentive in the form of inclusion in a draw for one of several gift cards worth up to \$100. From these first two Australian samples, a total of 611 students started the survey and 538 finished.

A third Australian sample was recruited several weeks later for the purposes of increasing sample size, diversifying the participants to other majors, and determining whether sampling method variation would produce different results. For this, an email invitation was sent to all of the university's 43,000 plus students. A total of 785 students started the survey; 685 finished. The three Australian samples jointly provided usable information about 1,086 Australian partnerships. The response rate for the first two Australian samples was 34% (611/1801), which is comparable to other on-line surveys administered at this university. Due to the nature of the sampling method, no response rate can be calculated for the third Australian sample nor the American sample.

The average age for the US sample was 21.6 (sd = 5.1). The average age for the three combined Australian samples was 25.8 (sd = 8.6). The 1,758 participants from both countries reported on 1,273 current partnerships and 723 former partnerships. Never married participants were 60.0% of the combined Australian sample and 85.1% of the US sample. Among the Australian participants, 15.9% were married, 15.9% were in common-law relationships, 1.9% were separated, 2.1% were divorced, and 4.2% were engaged. Among the US participants, 5% were married, 0.9% were in common-law relationships, .1% were separated, 1.8% were divorced, and 7.1% were engaged. In the US sample, 52.4% of respondents classified themselves as White, 20.2% Black, 21.6% Hispanic/Latino, 1.4% Asian, and 4.4% as another ethnic group. In the Australian sample, 83.0% of respondents classified themselves as White, 1.1% Aboriginal, .1% Torres Strait Islander, 6.7% Asian, and 9.1% as another ethnic group.

Lower male proportions in the US sample (22.1%) and the Australian samples (16.5%) were expected in part because the gender ratios of the various social science fields of study from which students were sampled are weighted toward females (Straus & Mickey, 2012); more specifically, there were approximately three females to one male in the majors/classes surveyed at both the Australian and US school. In addition, prior research has consistently found that female survey response rates are higher than male response rates (Porter & Whitcomb, 2005).

Procedure

Survey.

A computer-administered on-line survey containing questions adopted from Straus and colleagues' conflict tactics scale (CTS) was developed by the author. The survey retained the original CTS instructional preamble and seven of the original CTS items with wording nearly identical to the CTS, which measured both victimization and perpetration as reported by each respondent. The United States and Australian survey versions were identical other than minor terminology changes where necessary. For example, it is customary to ask US individuals to "check" a box, while in Australia it is customary to "tick" a box.

The main portion of the questions for the seven victimization questions were as follows: a) threw something at you that could or did hurt, b) threatened to hit you or throw something at you, c) slapped you, d) pushed or shoved you, e) used a gun or knife on you, f) punched, hit, or kicked you, and g) choked you.

The perpetration questions were modified to refer to the participant doing the act toward their partner. There were too few gun/knife incidents to permit analysis, so these were removed. This left six perpetration and six victimization questions for a total of twelve event types per relationship. Although event frequency questions with answer categories identical to the CTS were included in the instrument, the answer categories used for the current analysis were a

yes/no version of these items indicating whether each behavior ever occurred during the relationship. Although the current analyses do not combine these items into a traditional summative IPV scale, for comparisons with other studies (e.g., Straus & Mickey, 2012), the internal consistency reliability in the form of Chronbach's alpha was calculated as .86 for a victimization scale and .77 for a perpetration scale.

There were minor alterations of the survey's administration from the procedures commonly used in CTS-based IPV surveys. The ones most relevant are as follows. First, only respondents who indicated that they had engaged in at least one "serious relationship" since high school were administered the CTS portion of the survey. The term, "serious" was used after preliminary testing indicated that numerous university-aged respondents in both Australia and the United States suggested this term as the one most commonly used among their peers to describe a relationship having emotional attachment and involving more than one or two dates.

Second, rather than adhering to the traditional CTS method of asking respondents only about their current relationship, respondents who had both a current and past partner were given IPV questions inquiring about events with both the current and the most-recent past partner. Respondents who indicated that they had only a current partner were asked only about this one individual, while those who indicated that they had dated a serious partner but were not doing so currently were asked about their most-recent past partner. The survey inquired about two partners because administration methods that inquire only about current partners miss a substantial amount of violence during the high-risk times of separation and divorce (Johnson, 2008; Saunders, 2002; Taft, Hegarty, & Flood, 2001; Yllö, 1993) as well as violence that *causes* relationship dissolution (Ackerman, 2012). Other deviations from customary CTS administration methods were minor and not relevant to the current study. Details about these are described by Ackerman (2015).

Analytic Strategy

Multi-level analyses.

Multi-level/hierarchical models have been well-established in the psychological literature (Stephen W. Raudenbush, 2001; Stephen W. Raudenbush, Brennan, & Barnett, 1995; Stephen W. Raudenbush & Bryk, 2002; Stephen W. Raudenbush & Chan, 1993). Although unable to locate prior research in the psychological or criminological literatures that has applied a multi-level model to the study of IPV in a manner similar to that described below, research in the medical literature has utilized parallel multi-level analyses when comparing event endorsements with different IPV screening instruments (MacMillan et al., 2006). In addition, an example by Raudenbush (2003) using self-reported behavioral acts in a multilevel model (including IPV) is somewhat similar to the one employed in the current analyses. Because the nature of the present analyses permits the results to be interpreted similarly to traditional logistic regression, only a few of the most relevant methodological details are presented in this section.

An important detail to note is that these data and the purpose of the current analyses requires a multi-level technique due to the data's "nested" format, common when non-independent measurements are included for each research participant. In the present case, several participants had more than one romantic relationship, so relationships are "nested" within individuals. More importantly, the units of analysis are answers to individual CTS questions about IPV events (whether each event is over-reported or not), which are "nested" within relationships. In sum, there are three analytic levels: (1) IPV events, (2) relationships, and (3) individuals, and for this reason, multi-level models are required to provide appropriate standard error calculations for these data (Schwartz & Ackerman, 2001). Because the purpose of the present study is to model the predictors of over-reporting regardless of the analysis level rather

than proportion the variance in over-reporting among these levels, the fact that the number of relationships is not much larger than the number of participants is not particularly relevant.

Because the present study uses samples from different universities, these could be considered an additional analytic level. Preliminary modelling with this additional complexity did not result in substantial differences between models that used the fourth level and those that considered the sample type as a control variable. For this reason, the more parsimonious three-level hierarchical logistic regression model was used where the analyses regressed a dichotomous indicator of over-reporting on event-, relationship-, and person-level characteristics.

There is one final point to mention regarding the nature of the analytic units in the current study before moving ahead. The analyses presented herein use individual CTS items as the analytic unit *without regard to whether each was or was not endorsed* as having occurred. This produces information about the *types of people* more or less likely to over-report IPV rather than information that might predict *whether or not a particular IPV report was misreported*. Because the difference between these two is subtle and difficult to understand, an example is in order. Assume the hypothetical case involving 100 male and 100 female research participants who report on whether or not 6 different types of IPV events have transpired. The 100 hypothetical males endorse 10 of the possible 600 total events (100 males x 6 events each) of which 8 are found to be over-reports. In this case 80% of the male victimization endorsements are over-reports (8/10) while the male over-reporting *rate based on the percentage of total items in the denominator* is only 1.3% (8/600). Therefore, depending on which type of over-reporting figure is being described (and therefore which denominator is used), the over-reporting rate in this hypothetical example might be described as being either 1.3% or 80%.

To illustrate a gender comparison, perhaps the 100 females endorse 40 of the possible 600 total events of which 15 are found to be over-reports. In this case 38% of the *endorsements*

are over-reports (15/40, a *lower* percentage than 80% for the males), while the female over-reporting rate based on the percentage of total items in the denominator is 6.7% (15/600, a *higher* percentage than 1.3% for the males). So, in this hypothetical example, female endorsements are *less* likely than male endorsements to be over-reports, while at the same time, females are *more* likely than males to over-report when asked questions about each event type.

Variables

Overview.

Details of the three analytic levels are as follows:

Level 3: Individuals – $N = 1,758$. There were 1,758 respondents who provided the data necessary for the present analyses, 1,428 females and 330 males.

Level 2: Relationships – $N = 1,996$. Each participant was asked questions that determined whether she or he had current and/or past romantic partners. Some participants had current partners, some had past partners, and some had both. Because 238 participants reported on two opposite-sex partners, there are more relationships than participants. The multi-level model properly accounts for the non-independence of the two relationships per person.

Level 1: CTS events - $N = 23,952$. Each participant was asked several questions about victimization at the hands of their partner and corresponding questions about perpetration against their partner. Six victimization and six perpetration questions were included in the analysis for a total of twelve questions asked about IPV events. Because there were 1,996 relationships in the sample, the N at the event level (level 1) is $1,996 \times 12 = 23,952$. This represents the fact that each respondent had 12 *opportunities* to report (and thus over-report) a victimization or perpetration for each relationship regardless of whether or not they did so.

Dependent Variable.

Over-reporting. If a respondent endorsed one of the CTS questions about whether they were victimized by a partner or whether they perpetrated against a partner, the computer-administered survey next asked about the frequency with which the act occurred. The frequency question was followed by a “multiple-selection” question inquiring about event context. “Multiple-selection” refers to a question type where the respondent may select from one or more response-alternatives [not mutually exclusive] by checking appropriate boxes on the internet-based survey form. The context question for each victimization item began, “Which of the following apply to the most serious incident when [your partner] [event]?”, where [your partner] refers to either “Current Partner” or “Past Partner” and [event] refers to one of the CTS-type items. The context question contained eight response possibilities addressing the degree of injury, the intentionality of the act, whether the respondent considered the act to be self-defense, and other pertinent event details.

If the participant considered the event as self-defense, it was re-coded as not occurring for the purposes of the present study. Due to the complexity of self-defense as a separate form of over-reporting, a different manuscript will address self-defense endorsements.

The two responses relevant for the over-reporting measure are, “accident” and “joking, playful, or humorous – no one took it seriously.” If the respondent endorsed one of the CTS items and subsequently indicated that either (or both) of these two event details were applicable, the endorsement was considered an over-report. In other words, if a participant interpreted the questions in a very literal way, she or he may have reported acts such as slapping and hitting during initial questioning that an independent observer would not have classified as true partner violence. When asked follow-up questions about the context, however, a better indication of the event’s nature becomes apparent. Similar procedures were followed for classifying over-reports

for questions about respondent *perpetration* against both current and former partners using minor alterations to the question wording and response categories as appropriate.

Independent variables.

Individual level (level 3).

Gender/male. The 1,758 respondents were asked about their biological sex. This was coded 1 for male, 0 for female.

Age. This variable was included because age is associated with both maturity and experience in romantic partnerships. It was assumed that more mature and experienced individuals are less likely to over-report IPV because they have a better understanding of the seriousness of partner violence and are therefore less likely to mix accidents and playful behavior with purposeful and consequential aggressive acts. Age was grand-mean centered in order to provide the ability to better interpret the intercept of the regression model.

Relationship status. The respondents were coded as to whether they were never-married, currently-married, in common-law marriages (“defacto” relationships in Australia), separated, engaged, or divorced.

Country/sample. The respondents were coded to indicate from which sample they were drawn. The samples were labelled US, Australian 1, Australian 2, and Australian 3.

Relationship level (level 2).

Past/current partner. As noted above, some participants had current partners, some had past partners, and some had both. Data related to current partners were coded 1; data related to past partners 0. This variable was included to determine whether differences occur in over-reporting for current versus past partners for several reasons including the fact that participants may re-assess the meaning and interpretation of the event after a relationship ends (Hamby & Gray-Little, 2000).

Event level (level 1).

IPV event type. There were six victimization and six perpetration questions for each relationship. Model 2 included a dichotomous variable indicating whether the question referred to victimization or perpetration. In the full model (Model 3), a dummy variable representing each of the 12 event types was included, with the exception of being pushed by a partner. This was the event most likely to elicit over-reporting, so pushing-victimization was used as the reference.

Results

Table 1 illustrates the extent of over-reporting and provides basic gender comparisons. Of the 1,428 female participants who reported on 1,622 current or former relationships, 14.1% reported one or more victimizations by a partner and 13.9% reported one or more perpetrations against a partner. The 330 male participants reported on 374 current or former relationships where 19.4% of male participants reported one or more victimizations by a partner and 4.5% reported one or more perpetrations against a partner. These figures are roughly in line with prior studies based upon university-aged individuals, which often show that a larger proportion of males than females report being victimized (Straus & Mickey, 2012).

Because there were six victimization and six perpetration questions asked for each relationship, for the female sample there were 9,732 possible opportunities to endorse a victimization event and another 9,732 possible opportunities to endorse a perpetration event (1,622 x 6). The smaller male sample had 2,244 possible opportunities for each (374 x 6). As a preliminary step before the main analyses, the total number of victimization and perpetration events were calculated for each gender. Females endorsed 513 (5.3%) out of the possible 9,732 victimization events while the males endorsed 201 (9.0%). For the perpetration events, females endorsed 433 (4.4%) and males endorsed 27 (1.2%).

Of most interest in the Table 1 results is the events that were self-classified as over-reported. Overall, the male and female victimization and perpetration over-report totals were 259, which represents 1.1% of the opportunities the entire sample had to over-report either type of event. This figure also means that 22.1% of all item endorsements across both genders were over-reported – more than one in five.

The gender breakdown indicates that of the 513 female victimizations, 79 were labelled as either accidents or events that were, “joking, playful, or humorous – no one took it seriously.” These represent 15.4% of the 513 victimization endorsements and .8% of the 9,732 opportunities to endorse a victimization event. In contrast, males had 53 victimization over-reports. This number is 26.4% of the 201 male victimizations and 2.4% of their 2,244 opportunities to make an over-report. Together, the male and female figures indicate that the proportion of males who take the opportunity to over-report victimizations is three times that of females. At the same time, when a victimization item is endorsed by a male, it is almost 60% more likely to be an over-report than a victimization endorsement made by a female (26.4% versus 15.4%).

The over-reporting figures for perpetration are similarly problematic, but show a different pattern by gender. Females over-reported 112 perpetration events, which represents 1.2% of their opportunities and 25.9% of their total perpetration endorsements. Males over-reported 15 perpetration events, which represents .7% of their opportunities and 55.6% of their total perpetration endorsements.

In sum, males over-reported victimizations at a much higher rate than did females (2.4% versus .8%) while females over-reported perpetrations at a higher rate (1.2% versus .7%). If we look at only reported events, however, male endorsements for both victimizations and perpetrations were both more likely than were the female endorsements to be over-reports.

<Table 1 about here>

<Table 2 about here>

Tables 2 and 3 best address this study's main research questions, which focus on whether there are factors in addition to gender that predict IPV over-reporting and how consistent these are across various event types and samples. In other words, the final models examine whether there are statistical interactions between gender and many of the various factors included in the analyses. Because the full analytic model contains numerous statistical interactions, the analysis is separated into three models of increasing complexity. The three models permit easier interpretation, especially for those less familiar with multi-level analysis or multiplicative models. Population-average results for the multi-level logistic regressions are shown because these, rather than unit-specific results, are best to estimate population proportions (Neuhaus, Kalbfleisch, & Hauck, 1991).

Table 2 provides information quite similar to Table 1, although in a more detailed fashion. Table 2 can also determine whether Table 1's gender differences may result due to several compositional differences between the male and female samples (Kitagawa, 1955). For example, perhaps the males were younger than the females and younger individuals are more likely to over-report victimizations. If so, this might explain why Table 1 indicates that males over-reported victimizations to a greater degree than did females. In addition, Table 2 also permits an examination of how additional individual-level characteristics affect over-reporting.

With the exception of age, all variables are dichotomies. For this reason, along with grand-mean centering age and coding males as one and females as zero, the odds-ratio of the intercept of model 1 in Table 2 (.013) represents the expected rate of victimization over-reporting for females who are of the average age of the sampled respondents and coded as zero on all other variables. Converting Model 1's .013 intercept to a probability (probability = $\text{odds}/(1+\text{odds})$) gives .0128 (1.28% - quite close to the 0.8% figure from Table 1).

An important point to keep in mind is that this rate, while a seemingly small number, represents the *percentage of all event possibilities that are over-reported* and not the percentage of over-reported endorsement. Model 1 also indicates that after accounting for other variables, males were 40% more likely to over-report than females (OR = 1.400, $p = .017$) and over-reports become less likely as age increases (OR = .950, $p < .001$). A ten-year age increase is associated with a 40% decrease in the odds of over-reporting ($e^{-.5} = .60$).

Over-reporting did not differ across relationship status in any substantial way. There was a statistically significant difference in over-reporting between the first Australian sample and the US reference sample. Participants in this Australian sample were almost 46% less likely to over-report relative to the reference group (OR = .540, $p < .001$). There was no statistically significant difference between the rates of over-reporting for current versus former partners.

Model 2 adds a control indicating whether victimization or perpetration questions were addressed for each analytic unit, and a [cross-level] statistical interaction term that tests whether victimization events (versus perpetration events) are more likely over-reported by males or females. The male x victimization interaction coefficient is 1.443, the male coefficient is -.554, and the main effect of victimization is -.255, with an intercept of -4.173. Since the interpretation of these figures is quite difficult in their present form, it is helpful to convert these results into expected probabilities for over-reports for individuals of specific demographic groups. Note that in this context, expected probabilities are the same as expected proportions and will closely correspond to the proportions reported in Table 1. The 0/1 dichotomous coding of most variables and the centering of age permits the simplification of the regression equation for individuals of average age who are never-married and in the US sample to:

$$\text{Log-odds(over-reporting)} = -4.173 - .554(\text{male}) - .255(\text{victimization}) + 1.443(\text{male x victimization})$$

By exponentiating the right-hand side of the question and applying the formula:
 $\text{probability} = \text{odds} / (1 + \text{odds})$, the following probabilities can be calculated:

Over-reporting among males reporting about victimizations is 2.82%

Over-reporting among females reporting about victimizations is 1.18%

Over-reporting among males reporting about perpetrations is 0.88%

Over-reporting among females reporting about perpetrations is 1.52%

Again, we see that these figures are close to those from Table 1 (that do not contain statistical controls) where the most over-reporting occurs for males reporting about victimization (2.82% of victimization queries were over-reported), while the lowest degree of over-reporting occurs for male perpetration (.88%). As in Table 1, after the inclusion of several control variables, male still over-reported victimizations more than females and females still over-report more perpetrations than males. The fact that these figures are quite similar to Table 1's and that the gender differences remain after the inclusion of the control variables indicates that Table 1's gender differences are not explained by the compositional differences controlled for in Table 2.

Recall that these figures represent the probability/proportion of the total number of IPV events over-reported (6 victimization and 6 perpetration events per relationship) regardless of whether or not the event was endorsed, rather than the proportion of over-reports among events reported to have occurred. For this reason, although these probability calculations may seem small, they represent potentially large degrees of error when we consider their implications. For example, Table 1 illustrated that the seemingly small 1.1% overall rate of over-reports translated into a 22.1% rate of endorsed events being over-reported, a large over-reporting error.

<Table 3 about here>

Table 3 (Model 3) provides nuance to the male x victimization interaction by using dichotomous indicators (dummy variable coding) for each event-level CTS item rather than the

single victimization/perpetration contrast in Model 2. There are eleven victimization or perpetration acts compared to pushing victimization (the excluded reference category), each paired with a male interaction term. Together, each pair indicates the degree to which each CTS act is over-reported and whether males over-report each act to a different degree than do females.

First among several patterns is that some acts are less likely over-reported than others. For example, when a respondent reports choking a partner, she or he is less likely to be over-reporting than when reporting about other perpetration or victimization acts (the choke-perpetration coefficient, is the one of highest negative magnitude). This figure can be interpreted most readily for females, because the choke x gender interaction becomes zero as does the main effect of gender and so these drop from the equation. For this reason, the choke odds-ratio of .152 means that females are 15.2% as likely (84.8% less likely) to over-report choking perpetrations than they are to over-report pushing victimizations (the reference). The negative choke x male interaction implies that males may be even less likely than females to over-report choking, although this interaction term is not statistically significant, so this does not imply good evidence of an accurate estimate of the gender difference for the over-reporting of this act.

Second, two male x event *victimization* interactions are statistically significant at the $p < .01$ level (slapping and punching) and one (threatening) is statistically significant at the $p < .10$ level. All have *positive* coefficients indicating that males are more likely than females to over-report these events. Last, five out of the six *perpetration* events (other than choking) had significant *negative* interaction coefficients. These indicate that males were less likely to over-report these perpetrations against female partners than were females against male partners. Table 3 also added sample x male interaction terms to determine whether there were important gendered differences in over-reporting among the samples. Here we see that not only was over-reporting lower among respondents in the first Australian sample, but also that gender

differences in over-reporting occurred in the third Australian sample that were not readily apparent in Table 2 before the sample x male interaction term was added.

Discussion

This study examines over-reporting on surveys using the CTS or its adaptations. It examines whether male and female over-reporting differs across samples drawn from two countries or across samples within the same country, whether there are gendered differences in the over-reporting rates for different IPV events purportedly measured by the CTS, and whether other common demographic characteristics like participant age affect IPV over-reporting. The succinct answer to all three questions is, “yes.” Moreover, the analyses indicate that the manner of over-reporting, because it is not consistent across gender or age, nor across type of IPV event or the sample, is a potentially serious problem in IPV surveys that rely upon scales such as the CTS, at least among university students, and particularly among male participants. Before a more detailed discussion of these findings, however, a mention of the study’s limitations is in order.

Limitations

A potential limitation in this study is a reliance upon student samples. Although some might question the degree to which we might generalize the problems illustrated by these results since they were based upon student samples, we should consider how much more likely it may be for people of average education to misinterpret the conceptual meaning underlying survey questions than it is for more educated university students to do so.

Another potential issue is that over-reporting is determined by an event context contingency question in the present study. This can be problematic when participants are asked to contextualize their own antisocial behavior. Prior research, for example, has shown that perpetrators may see their own actions as legitimate, unavoidable, or lacking in any negative consequence (Baumeister, Stillwell, & Wotman, 1990) causing them to over-claim that acts

against partners were accidents or not taken seriously. This problem may be far less serious, however, for victimization reports where a victim's perceptions about a partner's intent may be more accurate than the reported perceptions of a perpetrator who may be providing post-hoc rationalizations for anti-social behavior.

Finally, this research examines only a few ways in which IPV may be over-reported. Another form of over-reporting involves acts of self-defense not examined in the current analysis (preliminary analysis indicated that these were not common) and there are other over-reporting possibilities left unexamined. For these reasons, over-reporting estimates are likely conservative.

Research Implications

The variation by which research participants interpret CTS-based measures is surely a result, at least in part, of the designers' intent to increase disclosure of sensitive events. One way in which scale designers increase disclosure is by avoiding terminology that might overly stigmatize aggressive acts in favor of terminology describing these acts as more normative facets of the inevitable conflict occurring within most romantic relationships. This strategy is justified based upon the well-documented problems that using value-laden terms such as "abuse" may cause (Hamby, 2014; Hamby & Gray-Little, 2000).

The avoidance of stigmatizing terminology like "abuse," "battering," or even "violence," however, may lead to misinterpretation, or what has been called "meaning shifts" (Gaskell, O'Muircheartaigh, & Wright, 1994). While it may be somewhat difficult for scholars to understand how respondents could misinterpret a carefully-worded question, theory within the survey design literature, supported by experiments and cognitive interviews, suggests that survey participants frequently interpret conceptual meaning within survey questions in quite unintended ways (Schwarz, Strack, Müller, & Chassein, 1988; Tourangeau, Rips, & Rasinski, 2000).

Of course, it remains unknown whether “meaning shifts” or similar caused the over-reporting in this research. Patterns presented in this and prior studies, however, seem to point mainly to *unintentional* mechanisms of this type. In a prior study, the strongest predictor of over-reports was whether the participant’s relationship involved mutual (bi-directional) aggression (Ackerman, 2015). On one hand, a likely explanation for the strength of the bi-directional aggression reports is that respondents inclined to interpret questions in a very literal way when reporting about their partner’s behavior are equally likely to use a very literal interpretation when reporting about the behavior they direct toward their partners. On the other hand, Kimmel (2002) has suggested that respondents may attempt to justify their aggressive behavior by over-reporting their partner’s aggression toward them (i.e., “I hit her because she hit me first”).

If respondents, however, were motivated to intentionally fabricate IPV victimizations, why would we see differences among types of IPV events? Would not intentional fabricators be just as likely to endorse acts of choking as they would pushing? Why would the over-reporting rates of male and female fabricators be similar for choking reports while much different for other types of IPV events? Instead, the patterns observed in the present research seem more likely if the mechanism underlying victimization over-reports were of an unintentional nature where descriptions of different IPV events elicit a differential susceptibility to misinterpretation. For example, it is easier to understand how a “slap” might be interpreted to include a light-hearted tap on the hand to chastise a dinner companion for reaching for a bite of dessert (Dobash, Dobash, Wilson, & Daly, 1992), than it is to imagine how choking someone was an accident.

An important matter that stands out in the basic results is the relatively low rate of male self-reported perpetration (4.5% for any perpetrations and 1.2% of the event possibilities), especially when compared to the rate by which females reported being victimized by males (14.1% for any victimizations and 5.3% of event possibilities). If the participating students’

partners were also students eligible for survey participation and all had a single partner, we would expect that the rate by which females report victimization to be nearly equal to the rate by which males report perpetration. Although the partners' student status was not measured in this study, prior research suggests that university students would generally be attracted to other university students (Kandel, Davies, & Baydar, 1990). Compounding the suspicious nature of the male self-reported perpetration rate is that when the males did report perpetrations, they subsequently reported over half (55.6%) to have been either accidents or events that neither partner took seriously.

Although impossible to definitely know from the current data, one must wonder whether these patterns suggest that different mechanisms underlie perpetration and victimization over-reports. Perhaps "meaning shifts" underlie a substantial amount of victimization over-reporting while a combination of "meaning shifts" and post-hoc rationalizations underlie a substantial portion of the over-reporting of perpetrations. Regardless of whether over-reporting explains some or all of the long-standing gender-symmetry controversy, however, gender differences in over-reporting clearly have implications for testing theories that rely upon accurate accounting of the gendered patterns of violence.

Regardless of the underlying mechanisms, if over-reporting was consistent across groups, samples, and events, misreporting would simply inflate estimated IPV prevalence and likely attenuate correlations between IPV scales and other measures of theoretical interest. If gender differences in over-reporting were not present or were more consistent across groups and IPV event types, corrections might be more easily made. The fact that there are so many factors that interacted with gender to produce varying rates of over-reporting among males and females, however, implies that corrections may be impossible and that the use of the CTS may be quite

problematic in cases where accurate estimates of gender differences in aggression are important for any of a variety of theoretical or practical reasons.

Clinical and Policy Implications

An important finding was that age is a robust predictor of over-reporting, with an effect-size not substantially altered by which variables were included in the models presented. Over-reporting reliably decreased across age, perhaps suggesting that as individuals mature, they become more likely to properly interpret the underlying intent of IPV questions. Alternatively, older individuals (or those from clinical samples) may be less likely to engage in the types of behaviors toward partners that might result in over-reporting even if they continue to interpret CTS questions in equally literal or inaccurate ways.

It is unknown exactly why there were differences between some of the Australian samples and the US psychology pool sample. Perhaps speculation is in order. The first Australian sample had an over-reporting rate significantly lower than that of the US sample used as the reference category. This sample also had the highest proportion of criminology and criminal justice students, who were very likely to have taken classes that touched on matters of legally-consequential forms of IPV. For this reason, these participants may have been more easily able to understand that the survey intended to measure partner violence, and therefore may have been more likely to properly interpret the intended conceptual meaning of the survey questions.

The males and females in the second Australian sample had over-reporting rates most similar to that of the US sample (the main effects were not significantly different than the US sample in any of the three models and the interactive effect was not significant in the third model). This second Australian sample had the most psychology majors. Perhaps psychology students from both countries had been exposed to classes where relationship problems involving more insensitive or immature behaviors rather than illegal assaults were discussed.

The third Australian sample was the only one with statistically significant gender differences in over-reporting, where males were more likely to over-report than the US reference group, while females were significantly less likely to over-report than the reference. The rate of male over-reporting in this sample far surpassed that of all other groups. It is difficult to speculate about the underlying cause of this pattern for this sample. What is known is that due to the university-wide nature of the sampling frame of this group, its participants would have been much less likely to have had exposure to any classes that covered any matters addressed by the social sciences. Perhaps unfamiliarity with concepts addressed by these classes has something to do with the cause. Unfortunately, the participant's major was not asked, so this possibility cannot be adequately addressed with the current data.

According to this study, CTS performance may differ greatly across samples, especially those involving participants of differing ages. For this reason, replications of this research should be conducted on older individuals and alternative samples to determine whether similar problems might be found among groups other than university-aged individuals from non-clinical samples.

The gendered patterns illustrated in this study point to validity problems that will not only affect IPV prevalence estimates, but have the potential to misinform our understanding of IPV's etiology – at least in circumstances where an accurate assessment of gender differences in aggression is important. Clearly, this is particularly important in research on matters of gender parity in IPV perpetration. Without additional research, however, it would be premature to conclude that prior findings, especially in the area of gender parity, based upon CTS analyses are incorrect. Rather than invalidating gender parity research based upon the CTS, perhaps the current results suggest that better distinctions be made between various forms of aggression that have greatly differing consequences (or perhaps no consequences) for those involved.

Overall, however, these matters support the contention of scholars such as Follingstad and Bush (2014), who argue for a critical need to develop a new “gold standard” of IPV measurement. The current study illustrates that part of this need is to determine how to minimize or detect IPV misreporting. One possibility is to re-word items using more specific terminology to minimize erroneous item endorsements. Another possibility is to encourage *more inclusive* endorsement of items intended to measure a wide range of acts and rely upon follow-up items similar to those used in the present analysis to enable analysts to better differentiate among alternative forms of ostensibly aggressive behavior that might exist on a continuum ranging from playful or insensitive all of the way up to deadly. This alternative would permit scholars to include in their studies only events relevant to their particular analyses.

Conclusion

Although the results of this study call into question the validity of the CTS scale, particularly when administered to university-aged individuals, it is important to emphasize that we should not conclude a need to abandon the use of surveys in IPV research. A conclusion of this sort is particularly problematic because quantitative survey data is often the driving force that organizes policy initiatives (Bart, Miller, Moran, & Stanko, 1989). Moreover, a great deal of evidence suggests that IPV victims generally prefer self-completed questionnaires over face-to-face questioning in clinical settings (MacMillan et al., 2006).

Rather than abandon survey research, a better alternative is to determine the best ways to minimize or correct for all forms of misreporting problems. As argued by DeKeseredy and Schwartz, “perfect surveys on violence against women are not possible, but good ones can and should be done” (1998 p.5).

References

- Ackerman, J. M. (2015). Over-Reporting Intimate Partner Violence in Australian Survey Research. *British Journal of Criminology*, 56(4), 646-667. doi:10.1093/bjc/azv066
- Ackerman, J. M., & Field, L. (2011). The gender asymmetric effect of intimate partner violence on relationship satisfaction. *Violence and Victims*, 26(6), 703-724.
- Bart, P. B., Miller, P. Y., Moran, E., & Stanko, E. A. (1989). Editors' introduction. *Gender and Society*, 3, 431-436.
- Baumeister, R. F., Stillwell, A., & Wotman, S. R. (1990). Victim and perpetrator accounts of interpersonal conflict: Autobiographical narratives about anger. *Journal of Personality and Social Psychology*, 59, 994-1005.
- DeKeseredy, W. S. (1995). Enhancing the quality of survey data on woman abuse: Examples from a national Canadian study. *Violence Against Women*, 1(2), 158-173.
- DeKeseredy, W. S., & Schwartz, M. D. (1998). *Measuring the Extent of Woman Abuse in Intimate Heterosexual Relationships: A Critique of the Conflict Tactics Scales*. Retrieved from
- Dobash, R. P., Dobash, R. E., Wilson, M., & Daly, M. (1992). The myth of sexual symmetry in marital violence. *Social Problems*, 39, 71-91.
- Follingstad, D. R., & Bush, H. M. (2014). Measurement of intimate partner violence: A model for developing the gold standard. *Psychology of Violence*, 4(4), 369-383.
- Gaskell, G. D., O'Muircheartaigh, C. A., & Wright, D. B. (1994). Survey questions about the frequency of vaguely defined events: The effects of response alternatives. *Public Opinion Quarterly*, 58, 241-254.
- Grych, J., & Hamby, S. L. (2014). Advancing the Measurement of Violence: Challenges and Opportunities. *Psychology of Violence*, 4(4), 363-368.

- Hamby, S. L. (2005). Measuring gender differences in partner violence: Implications from research on other forms of violent and socially undesirable behavior. *Sex Roles - Special Issue: Understanding Gender and Intimate Partner Violence*, 52, 725-742.
- Hamby, S. L. (2009). The gender debate about intimate partner violence: Solutions and dead ends. *Psychological Trauma: Theory, Research, Practice, and Policy*, 1(1), 24-34.
doi:10.1037/a0015066
- Hamby, S. L. (2014). Intimate partner and sexual violence research: Scientific progress, scientific challenges. *Trauma, Violence, & Abuse*, 15(3), 149-158.
doi:10.1177/1524838014520723
- Hamby, S. L., & Gray-Little, B. (2000). Labeling partner violence: When do victims differentiate among acts? *Violence and Victims*, 15(2), 173-186.
- Johnson, M. P. (1995). Patriarchal terrorism and common couple violence: Two forms of violence against women. *Journal of Marriage and Family*, 57, 283-294.
- Johnson, M. P. (2008). *A Typology of Domestic Violence: Intimate Terrorism, Violent Resistance, and Situational Couple Violence*. Boston: Northeastern University Press.
- Kandel, D. B., Davies, M., & Baydar, N. (1990). The creation of interpersonal contexts: Homophily in dyadic relationships in adolescence and young adulthood. In L. N. Robins & M. Rutter (Eds.), *Straight and Devious Pathways to Adulthood* (pp. 221-241). New York: Cambridge University Press.
- Kimmel, M. S. (2002). Gender symmetry in domestic violence. *Violence Against Women*, 8(11), 1332-1363.
- Kitagawa, E. M. (1955). Components of a difference between two rates. *American Statistical Association Journal*, 50, 1168-1194.

- Langhinrichsen-Rohling, J. (2005). Top 10 Greatest "Hits": Important Findings and Future Directions for Intimate Partner Violence Research. *Journal of Interpersonal Violence*, 20(1), 108-118.
- Lehrner, A., & Allen, N. E. (2014). Construct Validity of the Conflict Tactics Scales: A Mixed-Method Investigation of Women's Intimate Partner Violence. *Psychology of Violence*, 4(4), 477-490.
- MacMillan, H. L., Wathen, C. N., Jamieson, E., Boyle, M., McNutt, L.-A., Worster, A., . . . Webb, M. (2006). Approaches to screening for intimate partner violence in health care settings: A randomized trial. *Journal of the American Medical Association*, 296(5), 530-536.
- Margolin, G. (1987). The multiple forms of aggressiveness between marital partners: How do we identify them? *Journal of Marital and Family Therapy*, 13, 77-84.
- Neuhaus, J. M., Kalbfleisch, J. D., & Hauck, W. W. (1991). A comparison of cluster-specific and population-average approaches for analyzing correlated binary data. *International Statistical Review*, 59(1), 25-35. doi:10.2307/1403572
- Porter, S. R., & Whitcomb, M. E. (2005). Non-response in student surveys: The role of demographics, engagement, and personality. *Research in Higher Education*, 46(2), 127-152.
- Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*, 52, 501-525.
- Raudenbush, S. W., Brennan, R. T., & Barnett, R. C. (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology*, 9, 161-174.

- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models* (2nd. ed.). Thousand Oaks, CA: Sage.
- Raudenbush, S. W., & Chan, W.-S. (1993). Application of a hierarchical linear model to the study of adolescent deviance in an overlapping cohort design. *Journal of Consulting and Clinical Psychology, 61*, 941-951.
- Raudenbush, S. W., Johnson, C., & Sampson, R. J. (2003). A multivariate, multilevel Rasch model with application to self-reported criminal behavior *Sociological Methodology, 33*(1), 169-211.
- Saunders, D. G. (1986). When battered women use violence: Husband-abuse or self-defense? *Violence and Victims, 1*(1), 47-60.
- Saunders, D. G. (2002). Are physical assaults by wives and girlfriends a major social problem. *Violence Against Women, 8*(12), 1424-1448.
- Schwartz, J., & Ackerman, J. (2001). In search of a dependent variable: A comment on Avakame, 1998. *Criminology, 39*(4), 969-979.
- Schwarz, N., Strack, F., Müller, G., & Chassein, B. (1988). The range of response alternatives may determine the meaning of the question: Further evidence on informative functions of response alternatives. *Social Cognition, 6*(2), 107-117.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The conflict tactics (CT) scales. *Journal of Marriage and Family, 41*(1), 75-88.
- Straus, M. A., & Mickey, E. L. (2012). Reliability, validity, and prevalence of partner violence measured by the conflict tactics scales in male-dominant nations. *Aggression and Violent Behavior. doi:10.1016/j.avb.2012.06.004*
- Taft, A., Hegarty, K., & Flood, M. (2001). Are men and women equally violent to intimate partners? . *Australian and New Zealand Journal of Public Health, 25*(6), 498-500.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.

Yllö, K. A. (1993). Through a feminist lens: Gender, Power, and Violence. In R. J. Gelles & D. R. Loseke (Eds.), *Current Controversies in Family Violence*. Newbury Park, CA: Sage.

Table 1
IPV Event Endorsements and Over-reporting

Participants and Event Endorsement	N	Percentage 1 ^a	Percentage 2 ^b	Percentage 3 ^c
Research participants	1,758			
Relationships	1,996			
Participants reporting one or more victimizations	266	15.1%		
Participants reporting one or more perpetrations	213	12.1%		
IPV event endorsements	1,174		4.9%	100.0%
Over-reports	259		1.1%	22.1%

a) of the 1,758 participants

b) of 23,952 victimization and perpetration event possibilities (1,996 x 12)

c) of 1,174 victimization and perpetration event endorsements

Table 2
Predictors of IPV Over-reporting Among Opposite-sex Partners (Preliminary Models)^a

Predictor	Model 1			Model 2		
	b	OR	Std. Err.	b	OR	Std. Err.
(3) Person-level (N=1,758)						
Male respondent	.336	1.400	.141*	-.554	.575	.218*
Age (grand-mean centered)	-.051	.950	.012*	-.051	.950	.012*
Relationship Status (ref.=never married)						
Married	-.101	.904	.190	-.089	.915	.185
Common-law	.168	1.183	.148	.155	1.168	.144
Separated	.660	1.935	.434	.685	1.983	.420
Divorced	-.112	.893	.312	-.108	.897	.306
Engaged	.269	1.308	.191	.283	1.327	.187
Sample (ref.=US sample)						
Australian 1	-.616	.540	.141**	-.619	.539	.136**
Australian 2	-.137	.872	.177	-.131	.877	.172
Australian 3	-.135	.874	.148	-.138	.871	.144
(2) Relationship-level (N=1,996)						
Current partner (ref.=former partner)	-.083	.920	.084	-.080	.923	.082
(1) Event-level (N=23,952)						
Victimization (ref.=perpetration)				-.255	.775	.079**
Cross-Level Interactions						
Male x victimization				1.443	4.235	.205**
Intercept	-4.315	.013	.086**	-4.173	.015	.094**

^a Multi-level population-average model with robust standard errors

+ p < .10

* p < .05

** p < .01

Table 3
Predictors of IPV Over-reporting Among Opposite-sex Partners (Full Model)^a

Predictor	Model 3		
	b	OR	Std. Err.
(3) Person-level (N=1,758)			
Male respondent	.404	1.498	.249 ⁺
Age (grand-mean centered)	-.048	.953	-.013**
Relationship Status (ref.=never married)			
Married	-.190	.827	.205
Common-law	.195	1.215	.155
Separated	.671	1.956	.433
Divorced	-.153	.858	.321
Engaged	.233	1.262	.194
Sample (ref.=US sample)			
Australian 1	-.627	.534	.143**
Australian 1 x male	-.069	.933	.506
Australian 2	-.102	.903	.202
Australian 2 x male	-.479	.620	.364
Australian 3	-.409	.665	.150**
Australian 3 x male	1.090	2.974	.331**
(2) Relationship-level (N=1,996)			
Current Partner (ref.= former partner)	-.093	.911	.088
(1) Event-level (N=23,952)			
Victimization (ref.=push victimization)			
Threw	-.497	.608	.134**
Threw x male	-.117	.889	.296
Threat	-.624	.536	.145**
Threat x male	.456	1.578	.252 ⁺
Slap	-1.049	.350	.136**
Slap x male	.786	2.095	.234**
Punch	-1.049	.350	.149**
Punch x male	.682	1.977	.255**
Choke	-.947	.338	.147**
Choke x male	-.470	.625	.333
Perpetration			
Threw	-.624	.536	.149**
Threw x male	-.793	.452	.307**
Threat	-.073	.929	.128
Threat x male	-1.344	.261	.326**
Slap	-.152	.859	.129
Slap x male	-1.635	.195	.377**
Push	.068	1.070	.121
Push x male	-1.485	.226	.323**
Punch	-.152	.859	.135
Punch x male	-1.265	.282	.329**
Choke	-1.881	.152	.189**
Choke x male	-.534	.595	.490
Intercept	-3.783	.023	.120**

^a Multi-level population-average model with robust standard errors

⁺ p < .10

* p < .05

** p < .01