HHS Public Access

Author manuscript

J Fam Psychol. Author manuscript; available in PMC 2021 February 01.

Published in final edited form as:

J Fam Psychol. 2020 February ; 34(1): 101–111. doi:10.1037/fam0000560.

Self-Reporting DSM-5/ICD-11 Clinically Significant Intimate Partner Violence and Child Abuse: Convergent and Response Process Validity

Richard E. Heyman,
New York University

Jeffery D. Snarr,
The College at Brockport

Amy M. Smith Slep,
New York University

Katherine J. W. Baucom,
New York University

David J. Linkh
U.S. Air Force

Abstract

DSM-5 and ICD-11 (proposed) now include criteria for clinically significant (a) intimate partner violence (IPV) and neglect and (b) child abuse and neglect. However, existing measures of IPV and child abuse do not allow for assessment of established criteria. The current study examines the convergent and response process validity of the Family Maltreatment (FM) measure of clinically significant physical and psychological IPV and child abuse. Participants (N = 126) completed the FM via computer and measures of IPV (Revised Conflict Tactics Scale; Straus et al., 1996) and child abuse (Parent-Child Conflict Tactics Scale; Straus et al., 1998) via paper-and-pencil. Participants who endorsed acts of aggression on the FM completed an audio-recorded computerized interview recounting the two most severe incidents. Verbalized incidents (n = 138) were coded for clinically significant family maltreatment. Results largely supported the convergent validity of the FM. Agreement of acts endorsed on the FM with those endorsed on convergent measures was excellent for IPV and physical child abuse, yet poor for psychological child abuse. Further, in support of the response process validity of the FM, comparison with observer ratings of interviews indicated few “false positives” and no “false negatives” on the FM across the examined types of clinically significant IPV and child abuse. In summary, the FM is a promising measure for the assessment of clinically significant physical and psychological abuse as defined in the DSM-5 and ICD-11 (proposed).

Correspondence concerning this article should be addressed to: Richard E. Heyman, Ph.D., Family Translational Research Group, New York University, 137 East 25th Street, 6th Floor, New York, NY 10010; Richard.Heyman@NYU.edu.
Richard E. Heyman, Amy M. Smith Slep, Katherine J.W. Baucom, Family Translational Research Group, New York University; Jeffry D. Snarr, College at Brockport; David J. Linkh, U.S. Air Force. Katherine J.W. Baucom is now at the Department of Psychology, University of Utah.

Keywords

intimate partner violence; child maltreatment; physical abuse; psychological abuse; reliable criteria; measurement

Clinically significant[1] partner and child abuse and neglect — partner or child acts/omissions associated with present distress or with a significant increased risk of suffering, death, pain, disability, or an important loss of freedom (Heyman & Slep, 2019) — has received increasing attention, including expanded criteria sets in the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5; American Psychiatric Association, 2013) and proposed criteria sets under consideration for the International Classification of Diseases, 11th Edition (ICD-11; Heyman et al., 2017; Heyman, Slep, Erlanger, & Foran, 2013a; Slep, Heyman, & Malik, 2013). Family maltreatment (an umbrella term for abuse and neglect) comprises eight separate forms: (a) physical, (b) psychological, and (c) sexual intimate partner violence (IPV); (d) intimate partner neglect; (e) physical, (f) psychological, and (g) sexual child abuse; and (h) child neglect. The varied forms of abuse and neglect have widely documented impacts on health (Garcia- Moreno et al., 2006; Gilbert, Widom, Browne, Fergusson, & Webb, 2009; Lagdon, Armour, & Stringer, 2014; Vachon, Krueger, Rogosch, & Cicchetti, 2015). Partner and child abuse and neglect have extensive research literatures on prevalence, etiology, and treatment (Capaldi, Knoble, Shortt, & Kim, 2012; Dixon & Graham-Kevan 2011; Korbin & Krugman, 2014; Smith et al., 2017; World Health Organization, 2013); are associated with greater health care utilization and costs (Bonomi, Anderson, Rivara, & Thompson, 2009); and are factors in precipitating, exacerbating, and maintaining mental and behavioral disorders (Patel, 2007; Schonbrun & Whisman, 2010; Trevillion, Oram, Feder, Howard, 2012).

Reliable and valid criteria sets for use by clinicians in field settings have been developed and field tested for the eight forms of family maltreatment (Heyman & Slep, 2006; Slep & Heyman, 2006). These criteria sets include both (a) the presence of an aggressive act (or, in the case of neglect, an omission), and (b) an impact meeting the “harm criterion”: significant impact (e.g., injury, fear, impaired development) or high potential for impact (e.g., strangulation, trying to hit with a car, likely developmental impact). Key definitional elements are incorporated in multiple criteria sets (e.g., subcriteria for “more than inconsequential fear reaction”) to provide continuity across the various forms of maltreatment. In this paper, we will refer to IPV or child abuse meeting the criteria as “clinically significant” (CS). In the field trials of the criteria’s performance, reliability (agreement between decisions at field test sites and those of a master reviewer) was high (over 90%; κ = .80–1.00; Heyman & Slep, 2006; Slep & Heyman, 2006) and stayed above 90% even when disseminated worldwide (Heyman & Slep, 2009). This is in notable contrast

[1]“Clinically significant” is commonly used to denote problems that exceed diagnostic thresholds separating notable but subthreshold problems (e.g., periods of depressed mood lasting days) from those exceeding diagnostic thresholds (e.g., symptoms meeting the criteria for Major Depressive Disorder). Since DSM-IIIR, clinically significant problems required both (a) a signature symptom pattern and (b) harm (e.g., “present distress . . .or disability . . . or significantly increased risk of suffering death, pain, disability, or an important loss of freedom” harm (American Psychiatric Association, 1987, p. xxii). Although “clinically significant depression” does not appear in the DSM, it has been used nearly 9,000 times in the scientific literature. Likewise, “clinically significant intimate partner violence” does not appear in the DSM or ICD; we have employed this usage to distinguish “clinically significant” (i.e., diagnostic guideline- meeting abuse or neglect) from sub-threshold problems (i.e., a single push that does not meet the harm criterion for impact.)

to similar comparisons conducted for individual DSM disorders (meta-analytic mean κ =. 27; Rettew, Lynch, Achenbach, Dumenci, & Ivanov, 2009). The physical and psychological IPV and child abuse criteria sets have evidence of convergent validity, established through relations of child and partner physical abuse and psychological abuse with a variety of hypothesized outcomes such as relationship satisfaction, depressive symptoms, suicidality, and alcohol problems (Foran, Heyman, Slep, & US Air Force Family Advocacy Program [USAF FAP], 2011, 2014; Foran, Slep, & Heyman, 2011; Langhinrichsen-Rohling et al., 2011; Slep, Foran, Heyman, & Snarr, 2010; Slep, Foran, Heyman, Snarr, & USAF FAP, 2015; Wojda et al., 2017). Evidence of construct validity for the physical CS-IPV criteria can also be inferred from the hypothesized finding that women, compared with men, committed more IPV *acts* but were more frequently the victims of clinically significant physical and psychological IPV (Foran, Slep, Heyman, & USAF FAP, 2011).

This work led to the adoption (with modifications) of these criteria sets in the DSM-5 and ICD-11 (proposed).

Measuring Clinically Significant Family Maltreatment

There are well over 25 measures of physical, psychological, and/or sexual IPV (many reviewed in Centers for Disease Control and Prevention's [CDC] *Compendium of Tools*; Thompson, Basile, Hertz, & Sitterle, 2006). Most have some evidence for internal consistency and for various forms of construct validity (Thompson et al., 2006). Child maltreatment research has relied less on self-reports of general population families, although measures such as the Parent-Child Conflict Tactics Scale (CTSPC; Straus, Hamby, Finkelhor, Moore, & Runyan, 1998) and the ISPCAN Child Abuse Screening Tool—Parent Version (Runyan et al., 2009) have gained traction. However, existing measures of IPV and child maltreatment are not well-suited to operationalize clinically significant maltreatment, as their development predates CS criteria in military guidelines, DSM-5, and ICD-11 (proposed). Most IPV studies (e.g., Ellsberg & Heise, 2005; Straus & Gelles, 1990; Tjaden & Thoennes, 2000) solely report or emphasize prevalences of a single act of physical assault or purported psychological aggression even when impact was measured. Studies of child maltreatment using parental reports in general population families have almost always either measured acts alone (e.g., Straus et al., 1998; Runyan et al., 2009) or have avoided direct measurement of maltreatment altogether, relying instead on factors linked to child maltreatment potential (Milner, 1994), perhaps due to requirements to report families who admit to impactful physical or psychological acts. Clearly, measures of impactful, clinically significant family maltreatment are needed, and their validity must be tested and reported.

This paper describes the development and testing of the Family Maltreatment (FM) measure, which operationalized the clinically significant partner and child physical and psychological abuse criteria (a) tested and disseminated in the largest family protection agency in the U.S. (i.e., the military's Family Advocacy Program), (b) incorporated into the DSM-5, and (c) under consideration for the forthcoming ICD-11 (Heyman et al., 2018). The FM has been used in the largest series of surveys ever conducted regarding family maltreatment, U.S. Air Force-wide surveys in 2006, 2008, 2011, and 2014 comprising N>150,000 respondents (see Foran, et al., 2011, 2012, 2014; Langhinrichsen-Rohling et al., 2011; Lorber et al., 2017;

Slep et al., 2010, 2011a, 2011b, 2014, 2015; Snarr, Heyman, & Slep, 2010; Wojda et al., 2017).

Development of the FM

The FM was developed programmatically. In step 1, we conducted a systematic literature search and reviewed the questions and psychometric literature for all extant measures of physical and psychological IPV and child abuse and of child neglect[2] (available from the authors). Next, given that no measure conformed to the criteria for clinically significant IPV and child abuse, we drafted a measure corresponding to the criteria's elements that was capable of being administered in an anonymous, general population survey. In step 3, we conducted a content validity study (cf. Haynes, Richard, & Kubany, 1995) with 22 child maltreatment and 22 IPV experts of both the criteria (see Heyman & Slep, 2006, Study 1) and the FM. Their extensive feedback was used to refine both the criteria and measures. In step 4, we conducted six focus groups with partners/parents and pilot tested the FM with over 200 families. In step 5, the FM was administered to approximately 3,000 Air Force (AF) service members and their spouses as part of the 2003 biennial AF Community Assessment (CA).

The FM was developed with separate modules for physical CS-IPV (perpetration and victimization), psychological CS-IPV (victimization only), physical CS-child abuse (perpetration only) and psychological CS-child abuse (perpetration only). In this study, the full FM was administered, although it was designed to allow administration of as few as one of the modules, depending on administrators' interests. Furthermore, although the FM in this study (see Online Supplement 3) used skip patterns, it can also be administered with all act questions followed by the impact questions if any acts were endorsed (see Online Supplement 5).

Response Process Validity

Response process validity is one of four sources of construct validity identified in the *Standards for Educational and Psychological Testing* (American Education Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing, 2014). In the context of self-report measures, response process validity focuses on whether participants understand and respond to instructions and items as intended. In the case of the FM, although the items are face valid (e.g., "I spanked child on the bottom with a bare hand"), additional empirical testing should be done to ensure that the events reported fit within the definition of that act (Dillman, Smyth, & Christian, 2014). For example, if a

[2]Partner sexual abuse was not included as a target construct in the Air Force Community Assessment (because it was assessed in other surveys) and thus was not part of this development study. Suggestions for assessing clinically significant partner sexual abuse matching the DSM- 5/ICD-11 (proposed) criteria can be found in Heyman, Slep, Snarr, and Foran (2013). We attempted to develop a child neglect self-report. However, given that child neglect most often involves omissions — rather than commissions — and omissions (and their potential impacts that never come to pass) are, almost by definition, of low salience to the actor, pilot testing indicated that the self-report was of dubious validity. Thus, child neglect, too, was not included. Given the sensitivity of inquiring about child sexual abuse perpetration in a population self-report survey (and the low likelihood of self-reports), this construct was not assessed either.

respondent answers affirmatively to the item, “During the past year... using a stick, hairbrush, or some other hard object, I hit or spanked child,” but says during the interview that she spanked with a bare hand but never with an object, her response to that question would be inaccurate.

Current Study Aims

The purpose of the present study was to test the convergent and response process validity of the FM. First, we tested convergent validity by comparing participants’ FM questionnaire responses to those on other commonly used questionnaires of IPV and child abuse. Second, we tested response process validity by comparing respondents’ FM questionnaire responses to audio-recorded descriptions of aggressive incidents.

Method

The research was approved by both university (Stony Brook University) and AF Institutional Review Boards.

Participants

Participants were 126 AF active duty service members or spouses stationed at Sheppard Air Force Base (AFB) in Wichita Falls, TX (n = 65), or Ramstein Air Base and the surrounding military installations in Germany’s Kaiserslautern district (n = 61). Demographic characteristics were as follows: gender — male (67.5%), female (32.5%); relationship status — married (80.2%), single and never married (7.1%), legally separated or filed for divorce (3.2%), divorced or widowed (4.0%), not reported (5.6%); and number of children in the home — 0 (7.9%), 1 (21.4%), 2 (46.0%), 3 (16.7%), 4 or more (7.9%).

Participants were recruited from three sources. First, all families interviewed by the base’s child and partner maltreatment services agency (i.e., AF Family Advocacy Program) during the two-year period preceding recruitment were invited by Family Advocacy Program staff. (Voicemail messages were left if no one answered the phone.) Thirty individuals recruited via this source (23.8% of the total sample) completed the study.

Second, households with a married service member and at least one minor child were invited via phone by university research staff. (Voicemail messages were left if no one answered the phone.) Eighty-six individuals recruited via this source (68.3% of the total sample) completed the study.

Finally, at Sheppard AFB, Family Advocacy Program staff contacted first sergeants, told them about the study, provided available timeslots, and asked them to refer any service member volunteers who (a) would be eligible (i.e., were married and at least one minor child), and (b) could be spared from other duties for an hour. Ten service members recruited via this source (7.9% of the total sample) completed the study.

Procedure

Two university researchers, blind to the participants' recruitment source, collected the data over four consecutive days at each site. Participants came individually to the data collection site for a 45- to 60-min visit. Consent was obtained verbally and participants were provided with a hard copy of the informed consent form. However, to fully preserve participant anonymity, a waiver of documentation of written consent was granted by the Institutional Review Boards.

Participants were led to private rooms to complete the assessment. The FM and convergent validity measures were then administered with order of administration counterbalanced. If any IPV or parent-child aggression was reported on the FM, a computer- administered interview followed in which participants recounted the two most memorable incidents of IPV and two most memorable incidents of parent-child aggression in the last year. These interviews were audio-recorded for subsequent coding.

Following completion of the measures, the investigator answered participant questions, provided them with a list of community resources, and thanked them for their participation.

Measures

Family Maltreatment (FM) measure.—As noted above, the FM (see Online Supplement 3) comprises questions tailored to the criteria for physical and psychological CS-IPV and CS-child abuse. (See Online Supplement 2 for the FM and DSM-5/proposed ICD-11 maltreatment criteria and how the FM items operationalizes them.) In addition to reports on lists of specific acts of partner and parent-child aggression, respondents had the option of "writing-in" items (i.e., "Did something similar not listed here...") that were later coded as to whether they described an act.

Physical CS-IPV.—Participants were asked how frequently they and their partners had engaged in myriad aggressive behaviors in the past year. If at least one behavior was reported, participants were presented with a list of all endorsed behaviors and asked whether the incident(s) had resulted in any of a range of impacts (e.g., pain lasting at least 4 hours; bruise; fear). Physical CS-IPV was coded as present if (a) any act of physical aggression occurred, *and* (b) the act resulted in [1] injury; [2] significant fear; or [3] a high potential for more than insignificant physical injury (i.e., burning or scalding, using a weapon, or choking).

Psychological CS-IPV.—Only psychological CS-IPV *victimization* was assessed with the FM because perpetrators do not have access to the impact on victims' psychological well-being. Because of the ubiquity of psychologically aggressive acts without impact, participants were first screened for possible significant impacts. Participants were asked whether, in the past year, they had ever been so down, depressed, or stressed that it affected them almost every day for two weeks. If they answered yes, participants were then asked how much of their sadness, depression, and/or stress was related to things their partners said or did. Finally, participants were asked how frequently in the past 12 months they had feared for their own safety or that of someone they cared about due to their partners' behavior, and

whether fear of what their partners might do or say had interfered with their functioning (i.e., had prevented participants from pursuing work or educational goals, practicing their religion, obtaining health services, and/or contacting family and friends).

All participants who indicated having experienced at least some depression, stress, or fear as a result of their partners' behavior were asked how often (in the past year) their partners committed one of the listed behaviors. If at least one behavior was reported, participants were presented with each endorsed behavior and asked whether the behavior(s) had caused or contributed to the depression, stress, and/or fear they had reported earlier.

Psychological CS-IPV victimization was coded as positive if (a) at least one act of psychological aggression occurred in the past year that caused the participant to experience fear, stress, or depression, or (b) the participant feared for his/her own safety or that of someone s/he cared about due to the partner's behavior and this fear interfered with participant functioning.

Physical CS-child abuse.—During the focus group phase of FM development, parents consistently indicated that rather than reporting the *frequency* of acts of physical discipline, they would be more willing to disclose aggressive acts toward their children if they could qualify their behavior with the reason(s) they committed the acts. Thus, for each aggressive act (e.g., "I spanked the child on the bottom with a bare hand"), parents were presented with four possible options (for each of their children) to indicate their behavior/justification: "I did this to teach," "I did this to punish," "I did this because I was frustrated/lost my cool," and "I never did this."

Physical CS-child abuse was coded as positive if (a) any act of physical aggression occurred and resulted in at least one type of injury, or (b) no injury was reported, but at least one behavior was endorsed that had a high potential for injury (i.e., burning or scalding, beating, striking with an object, hitting with a fist, kicking hard, or choking).

Psychological CS-child abuse.—Parents reported on a range of psychologically aggressive acts toward their children. Possible responses for these relatively severe items ranged from "Never" to "More than once a day." Parents who endorsed (a) having engaged in an act of severe psychological aggression, and/or (b) having performed mild physical aggressive acts at least once or twice a day were presented with a list of the behaviors they had endorsed for each child and asked whether these parenting behaviors caused a significant impact.

Psychological CS-child abuse was coded as positive if at least one act (a) had a reported consequential psychological harm on the child (lower self-esteem, depression, nervousness/fear), (b) led to the child seeing a mental health professional or another helper, or (c) occurred at a frequency operationalized as causing high potential for more than inconsequential psychological harm (see Online Supplement 2).

Revised Conflict Tactics Scale (CTS2) and Parent-Child Conflict Tactics Scale (CTSPC).—Participants completed paper versions of the most widely used self-report

measures of IPV (CTS2; Straus, Hamby, Boney-McCoy, & Sugarman, 1996) and parent-child aggression (CTSPC; Straus, Hamby, Finkelhor, Moore, & Runyan, 1998). These scales ask respondents to indicate the frequency of their own (and, for the CTS2, their partners') psychologically and physically aggressive acts within the last year (e.g., swearing at, pushing, hitting). The CTSPC was completed for a single target child (i.e., the child the participant deemed to be the most difficult to raise).

Audio-recounted incidents.

Audio capture.—A sample of the computerized collection of audio-recounted incidents can be found in Online Supplement 4. The procedure was introduced as follows: "Lots of times, people filling out questionnaires complain that they can be misunderstood because they don't have the chance to explain their answers. This next activity will fix that. We will go back through some of your answers and ask you to talk into the microphone and describe some of your answers in more detail. When describing what happened, remember to include the beginning, middle, and end. Please give enough detail that someone listening would be able to picture exactly what happened as if she were watching a video of the event."

For the module for each form of IPV or child abuse, the computer displayed the following: "You said that during the past year you've used the following behaviors [when parenting/ during conflicts with your partner]," followed by the behaviors the respondent reported. Below the behaviors, the computer displayed the following: "and that, at least once, the following happened as a result, " followed by the impacts the respondent reported. Finally, the instructions concluded with the following: "Please think of the two most memorable times that these things happened in the past year. Think about where you and [your child/ your partner] were, what was going on, who else was around, and how the incident unfolded. Include the beginning, middle, and end." The respondent then pushed record and recounted the incidents. The procedure was repeated for each form of IPV or child abuse for which the respondent reported both acts and impacts.

Coding.—Respondents' accounts could be somewhat telegraphic, and the procedure did not allow for follow-up questions common in face-to-face interviews. Thus, each recorded incident was coded in two separate ways. In the non-blind condition, coders knew which acts and impacts the participants had endorsed on the FM; as they listened to each recording, they indicated which specific acts and impacts were mentioned. In the blind condition, coders were provided no information regarding which acts and/or impacts had been endorsed on the FM. Specific impacts were not coded; coders indicated only whether the impact criteria were met. Two clinical psychology Ph.D. students coded all audio recordings and were randomly assigned to the blind/non-blind condition for each recording.

Interrater reliability on the 138 verbalized incidents was excellent, with 96% agreement and Guilford's G = .92. (Because of the highly imbalanced cells, Cohen's **k** is extremely biased and G is the preferred statistic [Xu & Lorber, 2014]. As there are no published recommendations for the interpretation of G statistic values, we interpret G similar to Cohen's **k,** with values of .40-.59 "fair," .60-.74 "good," and values > .75 "excellent" [Cicchetti, 1994].) In 11 of the 15 disagreements, the non-blind rater indicated the incident

met criteria whereas the blind rater indicated it did not. These disagreements were likely due to the non-blind rater being able to make sense of a somewhat vague verbal description. Given the high concordance, we used the blind coding for comparisons, as it was the most conservative and independent.

Data Analysis

Convergent validity.—To test the convergent validity of the FM's maltreatment scores, we compared participants' reports on the FM to reports on the CTS2/CTSPC. As the CTS2 and CTSPC are not intended to match the clinically significant diagnostic criteria, only reports of aggressive acts were compared. We also examined disagreements to better understand discrepancies between reports on the two measures.

Response process validity.—To test the response process validity, we compared participants' reports on the FM with the blind ratings of the audio-recounted incidents. We also examined both "false positives" (i.e., cases where the FM indicated abuse but the audio-recounted incidents did not) and "false negatives" (i.e., cases where the FM indicated no abuse but the audio-recounted incidents did) to better understand discrepancies between measures.

Results

Convergent Validity of the FM

Table 1 presents yes/no agreement between the FM and the CTS2. Agreement ranged from good to excellent, with the exception of psychological parent-child aggression. These results, along with an in-depth examination of reasons for discrepancies, are summarized below by type of maltreatment.

Physical IPV (compared with CTS2).

(Comparison of FM and CTS2 physical IPV act items can be found in Supplemental Table 1a.) Agreement between the FM and the CTS2 was excellent on mild physical IPV perpetration and victimization (97% agreement on 116 and 127 ratings; G = .95 and .93, respectively) as well as severe physical IPV perpetration and victimization (97% and 93% agreement on 117 ratings of each; G = .93 and .86, respectively).

Disagreements on mild physical IPV occurred almost exclusively because participants reported a specific act of aggression on the CTS2 that was not on the FM (see Supplemental Tables 2a–b). Of the six reports of severe IPV perpetration and the 10 reports of severe IPV victimization, only two reports on each were made on both the FM and the CTS2. This may be explained by small wording differences (see Supplemental Tables 2c–d). For instance, five of six respondents who reported no severe acts of victimization on the CTS2 reported on the FM that "my partner punched or hit me." The CTS2 has an item "my partner punched or hit me with something that could hurt," whereas the FM has separate items for [a] punching/hitting and [b] striking with objects.

Psychological IPV (compared with CTS2).

(Comparison of FM and CTS2 psychological IPV act items can be found in Supplemental Table 1b.) FM skip patterns (requiring affirmation of possible impacts to triggering psychological aggression act questions) resulted in a lower number of ratings for psychological IPV compared with physical IPV. Despite this, agreement between measures was good (81% agreement on 31 ratings; G = .61).

Five disagreements were due to endorsement of partner insulting/swearing on the CTS2 but not the FM, and one disagreement was due to endorsement of "partner did something to spite me" on the CTS2 (no equivalent act on the FM; see Supplemental Table 2c).

Physical parent-child aggression (compared with CTSPC).

(Comparison of FM and CTSPC parent-child physical act items can be found in Supplemental Table 1c.) Agreement was excellent between measures on both mild (88% agreement on 97 ratings; G = .75) and severe physical parent-child aggression (97% agreement on 98 ratings; G = .94).

Of the 12 mild aggression cases where disagreement occurred (see Supplemental Table 2d), six parents endorsed an act ("grabbing") that was included in the FM but not the CTSPC; the remaining six parents endorsed at least one act on one measure that they did not endorse on the other. Of the three severe aggression cases where disagreement occurred (see Supplemental Table 2e), one parent endorsed an item ("threw/knocked child down") that was included in the CTSPC but not the FM; the other two indicated on the CTSPC that they had struck their child with an object somewhere other than on the bottom, but did not indicate where on the body the child was struck on the FM.

Psychological parent-child aggression (compared with CTSPC).

(Comparison of FM and CTSPC parent-child psychological act items can be found in Supplemental Table 1d.) Agreement between measures was poor on psychological parent-child aggression (47% agreement on 96 ratings; G = −.06, indicating less than chance agreement [Xu & Lorber, 2014]). However, this lack of agreement is entirely attributable to the inclusion of mild acts in the CTSPC (e.g., "shouted, yelled, or screamed at [child]," "threatened to spank or hit [child] but did not actually do it") that did not meet the threshold for inclusion in the FM (which, unlike the CTSPC, was designed to match criteria for clinically significant child abuse). The wide inclusiveness of the CTSPC can be detected in the prevalence of acts labeled psychologically aggressive (83%) compared with that of the FM (30%). See Supplemental Table 2e for detailed comparisons regarding reporting discrepancies.

Response Process Validity of the FM

Table 2 presents agreement on clinically significant maltreatment between the FM and ratings of audio-recounted incidents. Audio-recounted incidents were not provided for one-third (psychological CS-child abuse perpetration) to two-thirds (psychological CS-IPV victimization) of cases that met criteria for clinically significant maltreatment on the FM. Agreement between measures ranged from fair to perfect, with the exception of

psychological CS-IPV victimization. There were few cases where the FM indicated abuse but the audio-recounted incidents did not (“false positives”) and no cases where the FM indicated no abuse but the audio-recounted incidents did (“false negatives”). These results, along with an in-depth examination of potential “false positives,” are summarized below by type of maltreatment.

Physical CS-IPV.

Agreement between measures was excellent for both physical CS-IPV perpetration (86% agreement on 7 cases; G = .71) and victimization (100% agreement on 13 cases; G = 1.0). There was one potential “false positive” for physical CS-IPV perpetration. In that case, the audio-recounted incident did not mention the injury reported on the FM; the incident described did not have a high potential for more than inconsequential injury and was thus coded as not meeting criteria. This case should not necessarily be considered a false positive for purposes of FM validation, as it may reflect either (a) the paucity of verbalized information, or (b) an injury occurring during a different incident than the one(s) verbalized.

Psychological CS-IPV.

There was poor agreement between measures on psychological CS-IPV (29% agreement on the 7 cases; G = −.43). However, in all five potential “false positives,” participants reported on the FM experiencing a great deal of emotional distress as a result of their partners’ behavior over the past year; however, they did not mention impact in the audio recount. Thus, they did not directly contradict their FM responses in their interviews and should not be considered definitive “false positives” for purposes of response process validity.

Physical CS-child abuse.

There was fair agreement on physical CS-child abuse (77% agreement on 77 cases; G = .53). Nineteen participants verbalized a physical act toward their child that resulted in harm, or had more than reasonable potential to result in harm, whereas only one participant mentioned more than inconsequential injuries resulting from the act. The possible “false positives” resulted from a variety of scenarios. Several audio-recounted incidents ambiguously mentioned minor red marks that were not coded as above the injury threshold; when specifically queried on the FM, all met criteria. Five cases met criteria for child physical abuse on the FM only via “potential impact;” that is, they reported physically aggressive behaviors — in every case, striking the buttocks with flexible and/or hard objects — considered likely to cause injury. However, in the audio-recounted incidents, parents used the word “spank” to describe their behavior (not indicating whether the “spanking” was with a hand or an object). In all five cases, the coder judged “spanking” (implying use of a hand, not the object admitted to on the FM) to be unlikely to cause injury and coded the incident as not meeting criteria. Three cases met criteria on the FM via reports of welts; however, neither parent mentioned injuries in the audio-recounted incidents, and all used the word “spank” to describe their behavior. In sum, all FM/audio-recounted disagreements were due to ambiguous and/or minimized verbal accounts; none of the FM reports were directly contradicted by the verbal accounts.

Psychological CS-child abuse.

There was fair agreement between measures on psychological CS-child abuse (76% agreement on 41 cases; G = .51). No audio-recounted incidents were scored as meeting criteria because none of the 10 accounts mentioned qualifying negative impacts. In contrast, eight parents reported acts and considerable negative impacts on the child via the FM and two reported acts that exceeded the extreme frequency operationalization for meeting the "reasonable potential for psychological harm" impact criterion (see Online Supplement 2 for operationalization).

Discussion

DSM-5 and ICD-11 (proposed) borrow reliable and valid criteria (Heyman & Slep, 2006; 2009; 2019) operationalizing thresholds (based on the oft-employed "harm criterion" used throughout individual diagnoses) for couple and parent-child problems that cause pain, injury, an important loss of freedom, or death (see Foran et al., 2013; Heyman & Slep, 2019). Their development and inclusion in widely-used diagnostic systems recognizes the fact that researchers and clinicians routinely assess and treat individuals, couples, and families with relational problems, or individuals with disorders related to, or exacerbated by, relational problems. With the advent of such criteria, it becomes increasingly important to have tested measures that operationalize the criteria, including self-reports. (Structured clinical interviews were published in Heyman et al., 2013b and Slep, Heyman, Snarr, & Foran, 2013.) Such criteria and measures will lead to more optimal research communication (i.e., more standardized meanings of terms like "IPV" and "abuse"), improved assessment and screening, and better identification of and response to maltreatment in practice.

Research Implications

Given that established measures of IPV and child abuse pre-date diagnostic developments, they do not operationalize criteria for clinically significant abuse and neglect found in DSM-5 and ICD-11 (proposed). In this study, we examined the convergent and response process validity of one such measure, the Family Maltreatment (FM) measure. Overall, the results demonstrate convergent validity for the FM (as well as, conversely, for the CTS2 [Straus et al., 1996] and CTSPC [Straus et al., 1998]). Rates of agreement between the FM and the concordant CTS measures were generally high; where disagreements occurred, examination of data patterns suggest that respondents are sensitive to wording of acts and are not perfectly consistent in reporting acts. Further, the precise list of acts make a difference. Although the FM gives participants the option of writing in acts not found in the prior questions, relatively few did so. Respondents largely restrict their answers to the menu of items that developers provide; as a result, differences in these menus lead to differences in endorsements.

Several issues specific to particular maltreatment types are of note. First, psychological IPV is a challenging construct to measure because of the frequency of possible acts and their dependence on context (and often the interpretation of the victim). Because of this, more psychological acts are reported on the CTS2 than on the FM (whose acts also are drawn from the more severe end of the spectrum). Similar threshold and concomitant reporting

issues can be found with child physical and psychological aggression items on the CTSPC and FM.

This is the first study to examine response process validity of a family maltreatment measure. Response process validity is similar to "cognitive interviewing" (e.g., Dillman et al., 2014; Presser et al., 2004; Willis, 2005), whereby developers check to make sure that respondents are interpreting and responding to items in a manner consistent with the intention of the developers. To test this form of validity, we compared FM responses to orally recounted descriptions/explanations of the incidents. Independent ratings of these audio recordings, made by trained coders, indicated no "false negatives" on the FM across the examined types of CS-IPV (psychological CS-IPV perpetration, physical CS-IPV perpetration and victimization) and CS- child abuse (perpetration of psychological and physical CS-child abuse). There was high concordance between questionnaires' CS-IPV diagnoses and verbalized accounts of the incident for both physical (perpetration as well as victimization) and psychological CS-IPV. For both forms of CS-child abuse, all disagreements between the FM and the audio-recounted incident appeared to be due to ambiguous and/or minimized verbal accounts. This implies that although anonymous, computer-recorded vocalized accounts of incidents were employed as the only viable way to elicit an honest response about child abuse — given that identifiable accounts of child abuse would necessitate both reporting to child welfare authorities and warning participants that responding affirmatively would instigate such reporting — the inability to ask follow-up questions was particularly limiting for checking the validity of child abuse self-reports.

The FM was administered via computer, although it could be administered via paper-and-pencil (See Online Supplement 6). Computerized IPV screening has several advantages over other methods (at least in healthcare settings): (a) it is efficient and preferred by patients (e.g., Renker, 2008) and, (b) in a study of women in emergency departments, it resulted in higher screening, detection, and referrals compared with usual care (e.g., Trautman, McCarthy, Miller, Campbell, & Kelen, 2007). On the other hand, paper is simpler to administer and does not suffer from technological hurdles and glitches.

Finally, although the FM may be suitable for some research and clinical contexts, the modules may be too lengthy for use in medical or mental health practices or on general surveys. Heyman, Xu, et al. (2018) and Slep et al. (2018) described the development and testing of screeners for CS-IPV and CS-child abuse that (a) are maximally brief while still achieving high sensitivity and specificity for identifying clinically significant maltreatment; (b) assess (where pertinent) perpetration and victimization when either men or women are reporting; and (c) use act plus impact criteria as the reference standard. These authors recommended two-stage screening, whereby those screening positive are interviewed to determine if they meet the full criteria for a specific type of abuse or neglect.

Clinical and Policy Implications

Given the costs associated with family maltreatment, the development of reliable and valid criteria for clinically significant relational problems represented a crucial step for improving clinical practice and informing public policy. Measures such as the FM — that directly assess clinically significant relational problems and match DSM and ICD criteria — fill a

gap. Such measures provide clinicians with guidance about screening, assessment, diagnosis/classification, and intervention. Further, they provide an evidence-based foundation for public policy tied to these criteria.

Harmonized Assessment of Partner and Child Physical and Psychological Abuse

As noted in the introduction, the criteria for partner and child abuse and neglect share (a) theoretical thresholds (the notion of a "harm criterion" requiring an act plus clinically significant impact or high potential for impact) and (b) elements operationalized in identical or highly similar ways across types of abuse and neglect. By design, the FM has these communalities.

The FM modules for physical and psychological CS-IPV and CS-child abuse can be used separately or together. There are substantial reasons for measuring both partner and child abuse together. First, acts of parent-child and intimate partner aggression frequently co-occur, overlapping in 30–60% of U.S. families (Appel and Holden, 1998). Slep and O'Leary (2005), using a representative sample of families with children between 3 and 7 years of age from exurban New York, found 5% of homes reported experiencing severe acts of both IPV and parent-child aggression, almost certainly meeting the current criteria for physical CS-IPV and CS-child abuse.

Partner and child maltreatment are linked likely due to (a) person-based risk factors that are not specific to either the partner or parent-child dyad (e.g., personality characteristics, biological characteristics, psychological functioning, historical risk factors) and to (b) shared environmental or contextual risk factors (e.g., Jouriles, McDonald, Slep, Heyman, & Garrido, 2008; Knickerbocker, Heyman, Slep, Jouriles, & McDonald, 2007). Although some of the research by our group has examined common risk factors for CS-IPV and CS-Child Abuse (Wojda et al., 2017), most of the literature has examined either acts of potential abuse (for example, using the CTS or CTSPC) or has used indisputably clinically significant samples (from women's shelters or Child Protective Services) that are inappropriate to examine population-level risk. Thus, further work on etiology or interventions using clinically significant maltreatment is needed to better understand the phenomena.

Such work is not easy. Unless steps are taken to assess families anonymously, researchers are in a bind, as they are mandated reporters and would have to notify child protective services (CPS) if families reported CS-child maltreatment. Further, without anonymity, researchers would have to warn participants that positive reports would result in CPS notification, likely eliminating any true positive reporting. Although researchers can take steps to maintain anonymity throughout research studies (while meeting ethical obligations by notifying participants that, although their research reports are anonymous, anything they verbalize to investigators may result in CPS notification), clinicians do not have this option, as eliciting reports of CS-child maltreatment would automatically trigger reporting. Although clinicians may be aware of the overlap of forms of CS-family maltreatment and may strive to protect all vulnerable members of a family, they may be concerned about subjecting clients newly engaged in treatment to an investigation that parents typically perceive as intrusive rather than helpful and these concerns extend to (a) damaging the therapeutic alliance and (b) creating or increasing resistance and risk for dropout. Second, in

the vast majority of cases, concurrent CPS and IPV referrals do not lead to recommendations for services that differ from those that the family may already be receiving, so the benefit to the family is not always apparent. Thus, although clinical assessment of all forms of maltreatment can likely lead to better intervention planning, gathering such information without a clear, cogent, client-focused pre-assessment rationale may do more harm than good.

Limitations

First, studying response process validity without the capacity to ask follow-up questions was limiting. Studies that use anonymizing procedures and still allow for follow-up questions would eliminate some of the ambiguity that hampered the results of some forms of maltreatment. Second, a similar study that includes child and partner CS-sexual abuse is needed. (Reporting about CS-child neglect via questionnaires appeared to be impossible, given the low salience of parental omissions that don't result in actual harm. Structured clinical interviews are available and are likely the only viable mode of assessment.) Third, this sample comprised active duty AF members or their spouses. Although nothing about this study was military-specific, it's possible that data collection context (including the perceived negative career consequences of reporting maltreatment if anonymity assurances were not fully believed) could have affected results. Replications and extensions of this work in myriad civilian subpopulations are needed. Fourth, all self-report measures, whether collected via computer or paper-and-pencil, are limited by reporting biases. Finally, the sensitivity and specificity of both the CTS2/CTSPC and the FM should be examined versus a "gold standard" (e.g., diagnostic interview) in future studies. The current study examined convergence of the two measures and the match between the FM and verbalized accounts, but neither are sufficient to adequately examine sensitivity and specificity.

Conclusions

The FM has indications of convergent and response process validity and is a viable means of collecting self-reports of clinically significant partner and child physical and psychological abuse that conform to DSM5 and ICD-11 (proposed) criteria. Screeners derived from the FM (Heyman, Xu, et al., 2018; Slep et al., 2018) are also available when administration of the full FM is not viable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The opinions expressed in this article are solely those of the authors and do not necessarily represent the official views of the U.S. Government, the U.S. Department of Agriculture, or the U.S. Department of the Air Force. This research was supported by the United States Department of Agriculture (contract number CR-19191A545810, co-PIs Heyman and Slep).

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). Standards for educational and psychological testing (6th ed). Washington, DC: American Educational Research Association.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed). Arlington, VA: American Psychiatric Association.

Bonomi AE, Anderson ML, Rivara FP, & Thompson RS (2009). Health care utilization and costs associated with physical and nonphysical-only intimate partner violence. Health Services Research, 44, 1052–1067. doi: 10.1111/j.1475-6773.2009.00955.x [PubMed: 19674432]

Briggs EC, Thompson R, Ostrowski S, & Lekwauwa R (2011). Psychological, health, behavioral, and economic impact of child maltreatment In White JW, Koss MP, & Kazdin AE (Eds.), Violence against women and children, Vol 1: Mapping the terrain, pp. 77–97. Washington, DC: American Psychological Association. doi:10.1037/12307-004

Capaldi DM, Knoble NB, Shortt JW, & Kim HK (2012). A systematic review of risk factors for intimate partner violence. Partner Abuse, 3, 231–280. doi:10.1891/1946-6560.3.2.231 [PubMed: 22754606]

Cicchetti DV (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6, 284–290. doi: 10.1037/1040-3590.6.4.284

Dillman DA, Smyth JD, & Christian LM (2014). Internet, phone, mail, and mixed-mode surveys: The tailored design method. Hoboken, NJ: John Wiley & Sons.

Dixon L, & Graham-Kevan N (2011). Understanding the nature and etiology of intimate partner violence and implications for practice and policy. Clinical Psychology Review, 31, 1145–1155. doi: 10.1016/j.cpr.2011.07.001 [PubMed: 21851805]

Ellsberg M & Heise L (2005). Researching violence against women: A practical guide for researchers and activists. Washington, DC: World Health Organization and Program for Appropriate Technology in Health.

Foran HM, Beach SRH, Slep AMS, Heyman RE, & Wamboldt MZ (Eds.). (2013). Family problems andfamily violence: Reliable assessment and the ICD-11. New York, NY: Springer.

Foran HM, Heyman RE, Slep AMS, & US Air Force Family Advocacy Program. (2011). Hazardous drinking and military community functioning: Identifying mediating risk factors. Journal of Consulting and Clinical Psychology, 79, 521–532. doi:10.1037/a0024110 [PubMed: 21707135]

Foran HM, Heyman RE, Slep AMS & Air US Force Family Advocacy Program. (2014). Emotional abuse and its unique ecological correlates among military personnel and spouses. Psychology of Violence, 4, 128–142. doi:10.1037/a003453

Foran HM, Slep AMS, & Heyman RE (2011). Hazardous alcohol use among active duty Air Force personnel: Identifying unique risk and protective factors. Psychology of Addictive Behaviors, 25, 28–40. doi:10.1037/a0020748 [PubMed: 21244121]

Foran HM, Slep AMS, Heyman RE, & US Air Force Family Advocacy Program (2011). Prevalences of intimate partner violence in a representative U.S. Air Force sample. Journal of Consulting and Clinical Psychology, 79, 391–397. doi:10.1037/a0022962 [PubMed: 21480693]

Garcia-Moreno C, Jansen HAFM, Ellsberg M, Heise L, Watts CH, & WHO Multi-Country Study on Women's Health and Domestic Violence Against Women Study Team. (2006). Prevalence of intimate partner violence: Findings from the WHO multi-country bstudy on women's health and domestic violence. The Lancet, 368, 1260–1269. doi :10.1016/S0140-6736(06)69523-8

Gilbert R, Widom CS, Browne K, Fergusson D, & Webb E (2009). Burden and consequences of child maltreatment in high-income countries. The Lancet, 373, 68–81. doi :10.1016/ S0140-6736(08)61706-7

Haynes SN, Richard DCS, & Kubany ES (1995). Content validity in psychological assessment: A functional approach to concepts and methods. Psychological Assessment, 7, 238–247. doi: 10.1037/1040-3590.7.3.238

Heise L, & Garcia-Moreno C (2002). Violence by intimate partners In Krug EG, Dahlberg LL, Mercy JA, Zwi AB, & Lozano R (Eds.). World report on violence and health (pp. 87–121). Geneva, Switzerland: World Health Organization.

Heyman RE, Kogan CS, Foran HM, Burns SC, Slep AMS, ... Reed GM. (2018). A case-controlled field study evaluating ICD-11 proposals for relational problems and intimate partner violence. International Journal of Clinical and Health Psychology, 18, 113–123. doi: 10.1016/j.ijchp.2018.03.001 [PubMed: 30487916]

Heyman RE, & Slep AMS (2006). Creating and field-testing diagnostic criteria for partner and child maltreatment. Journal of Family Psychology, 20, 397–408. doi:10.1037/0893-3200.20.3.397 [PubMed: 16937996]

Heyman RE & Slep AMS (2019). Relational diagnoses and beyond In Friese B (Ed.), APA Handbook of Contemporary Family Psychology (pp. 19–34). Washington, DC: American Psychological Association Press.

Heyman RE, Slep AMS, Erlanger ACE, & Foran HF (2013a). Intimate partner maltreatment: Definitions, prevalence, and implications for diagnosis In Foran HM, Beach SRH, Slep AMS, Heyman RE, & Wamboldt MZ (Eds.), Family problems and family violence: Reliable assessment and the ICD-11 (p. 1–14). New York, NY : Springer.

Heyman RE, Slep AMS, Snarr JD, & Foran HM (2013b). Practical tools for assessing partner maltreatment in clinical practice and public health settings In Foran HF, Beach SRH, Slep AMS, Heyman RE, & Wamboldt MZ (Eds.), Family Problems and Family Violence: Reliable Assessment and the ICD-11 (p. 43–70). New York, NY: Springer.

Heyman RE, Xu S, Baucom KJW, Slep AMS, Snarr JD, ... U.S. Air Force Family Advocacy Program (2018). High sensitivity and specificity screening for clinically significant intimate partner violence. Manuscript submitted for publication.

Jouriles EN, McDonald R, Slep AMS, Heyman RE, & Garrido E (2008). Child abuse in the context of domestic violence: Prevalence, explanations and practice implications. Violence and Victims, 23, 221–235. doi: 10.1891/0886-6708.23.2.221 [PubMed: 18624091]

Knickerbocker R, Heyman RE, Slep AMS, Jouriles EN, & McDonald R (2007). Co-occurrence of child and partner maltreatment: Definitions, prevalence, theory, and implications for assessment, treatment, and policy. European Psychologist, 12, 36–44.

Korbin JE, & Krugman RD (2014). Handbook of child maltreatment. New York, NY : Springer

Lagdon S, Armour C, & Stringer M (2014). Adult experience of mental health outcomes as a result of intimate partner violence victimisation: A systematic review. European Journal of Psychotraumatology, 5, 24794. doi:10.3402/ejpt.v5.24794

Langhinrichsen-Rohling J, Snarr JD, Slep AMS, Heyman RE, Foran HM, & USAF Family Advocacy Program. (2011). Risk for suicidal ideation in the U.S. Air Force: An ecological perspective. Journal of Consulting and Clinical Psychology, 79, 600–612. doi:10.1037/a0024631 [PubMed: 21787046]

Lorber MF, Heyman RE, & Slep AMS (2017). A longitudinal investigation of the psychological health of United States Air Force base communities. Journal of Community Psychology, 45, 1033–1049. doi:10.1002/jcop.21909

Milner JS (1994). Assessing physical child abuse risk: The child abuse potential inventory. Clinical Psychology Review, 14, 547–583. doi:10.1016/0272-7358(94)90017-5

Patel V (2007). Mental health in low- and middle-income countries. British Medical Bulletin, 81–82, 81–96. doi:10.1093/bmb/ldm010

Presser S, Rothgeb JM, Couper MP, Lessler JT, Martin E, Martin J & Singer E (2004). Methods for testing and evaluating survey questionnaires. Hoboken, NJ: John Wiley and Sons.

Renker PR (2008). Breaking the barriers: The promise of computer-assisted screening for intimate partner violence. Journal of Midwifery & Women's Health, 53, 496–503. doi:10.1016/j.jmwh.2008.07.017

Rettew DC, Lynch AD, Achenbach TM, Dumenci L, & Ivanova MY (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. International Journal of Methods in Psychiatric Research, 18, 169–184. doi:10.1002/mpr.289 [PubMed: 19701924]

Runyan DK, Dunne MP, Zolotor AJ, Madrid B, Jain D, ... Youssef RM (2009). The development and piloting of the ISPCAN Child Abuse Screening Tool—Parent version (ICAST-P). Child Abuse & Neglect, 33, 826–832. doi:10.1016/j.chiabu.2009.09.006 [PubMed: 19854511]

Schonbrun YC, & Whisman MA (2010). Marital distress and mental health care service utilization among married adults. Journal of Consulting and Clinical Psychology, 78, 732–736. doi: 10.1037/a0019711 [PubMed: 20873908]

Slep AMS & Heyman RE (2006). Creating and field-testing child maltreatment definitions: Improving the reliability of substantiation determinations. Child Maltreatment, 11, 397–408. doi: 10.1177/1077559506288878

Slep AMS, Foran HM, & Heyman RE (2014). An ecological model of intimate partner violence perpetration at different levels of severity. Journal of Family Psychology, 28, 470–482. doi: 10.1037/a0037316 [PubMed: 25000132]

Slep AMS, Foran HM, Heyman RE, & Snarr JD (2010). Unique risk and protective factors for partner aggression in a large scale Air Force survey. Journal of Community Health, 35, 375–383. doi: 10.1007/s10900-010-9264-3xc [PubMed: 20373136]

Slep AMS, Foran HM, Heyman RE, Snarr JD, & Air US Force Family Advocacy Program. (2011). Risk factors for clinically significant intimate partner violence among active-duty members. Journal of Marriage and Family, 73, 486–501. doi: 10.1111/j.1741-3737.2010.00820.x

Slep AMS, Foran HM, Heyman RE, Snarr JD, & Air US Force Family Advocacy Program. (2015). Identifying unique and shared risk factors for intimate partner violence and clinically significant intimate partner violence. Aggressive Behavior, 41, 227–241. doi:10.1002/ab.21565 [PubMed: 27541201]

Slep AMS, Heyman RE, & Malik J (2013). Child maltreatment: Definitions, prevalence, and implications for diagnosis In Foran HF, Beach SRH, Slep AMS, Heyman RE, & Wamboldt MZ (Eds.), Family problems andfamily violence: Reliable assessment and the ICD-11 (p. 111–130). New York: Springer.

Slep AMS, Heyman RE, Snarr JD, & Air US Force Family Advocacy Program. (2011). Child emotional aggression and abuse: Definitions and prevalence. Child Abuse & Neglect, 35, 783–796. doi:10.1016/j.chiabu.2011.07.002 [PubMed: 22018518]

Slep AMS, Heyman RE, Xu S, Snarr JD, Foran HM, ... U.S. Air Force Family Advocacy Program (2018). Screening for clinically significant intimate child abuse. Manuscript submitted for publication.

Smith SG, Chen J, Basile KC, Gilbert LK, Merrick MT, ... Jain A. (2017). The National Intimate Partner and Sexual Violence Survey (NISVS): 2010–2012 state report Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Snarr JD, Heyman RE, & Slep AMS (2010). Recent suicidal ideation and suicide attempts in a large-scale survey of the U. S. Air Force: Prevalences and demographic risk factors. Suicide and Life-Threatening Behavior, 40, 544–552. doi:10.1521/suli.2010.40.6.54 [PubMed: 21198323]

Straus MA & Gelles RJ (Eds.) (1990). Physical violence in American families: Risk factors and adaptation to violence in 8,145 families. New Brunswick, NJ: Transaction Publishing

Straus MA, Hamby SL, Boney-McCoy S, & Sugarman DB (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. Journal of Family Issues, 17, 283–316. doi:10.1177/019251396017003001

Straus MA, Hamby SL, Finkelhor D, Moore DW, & Runyan D (1998). Identification of child maltreatment with the Parent-Child Conflict Tactics Scales: Development and psychometric data for a national sample of American parents. Child Abuse & Neglect, 22, 249–270. doi:10.1016/S0145-2134(97)00174-9 [PubMed: 9589178]

Thompson MP, Basile KC, Hertz MF, & Sitterle D (2006). Measuring intimate partner violence victimization and perpetration: A compendium of assessment tools. Atlanta, GA: Centers for Disease Control, National Center for Injury Prevention and Control.

Tjaden P, & Thoennes N (2000). Full report of the prevalence, incidence, and consequences of violence against women: Findings from the National Violence against Women Survey. Washington, DC: U.S. Department of Justice.

Trautman DE, McCarthy ML, Miller N, Campbell JC, & Kelen GD (2007). Intimate partner violence and emergency department screening: Computerized screening versus usual care. Annals of Emergency Medicine, 49, 526–534. doi:10.1016/j.annemergmed.2006.11.022 [PubMed: 17276547]

Trevillion K, Oram S, Feder G, & Howard LM (2012). Experiences of domestic violence and mental disorders: A systematic review and meta-analysis. PLoS ONE, 7, e51740. doi:10.1371/journal.pone.0051740

Vachon DD, Krueger RF, Rogosch FA, & Cicchetti D (2015). Assessment of the harmful psychiatric and behavioral effects of different forms of child maltreatment. JAMA Psychiatry, 72, 1135–1138. doi:10.1001/jamapsychiatry.2015.1792 [PubMed: 26465073]

Veroff J, Kulka RA, & Douvan E (1981). Mental health in America: Patterns of help- seeking from 1957–1976. New York, NY: Basic Books.

Willis G (2005). Cognitive interviewing: A tool for improving questionnaire design. Thousand Oaks, CA: Sage Publications, Inc.

Wojda AK, Heyman RE, Slep AMS, Foran HM, Snarr JD, & Air US Force Mental Health Division. (2017). Family violence, suicidality, and substance abuse in active duty military families: An ecological perspective. Military Behavioral Health, 5, 300–312. doi:10.1080/21635781.2017.1343698

World Health Organization. (2013). Responding to intimate partner violence and sexual violence against women: WHO clinical and policy guidelines. Geneva: World Health Organization

Xu S, & Lorber MF (2014). Interrater agreement statistics with skewed data: Evaluation of alternatives to Cohen's kappa. Journal of Consulting and Clinical Psychology, 82, 1219–1227. doi:10.1037/a0037489 [PubMed: 25090041]

Table 1.

Comparison of acts endorsed on FM and on CTS2/CTSPC[a]

		FM (Acts only)		G
CTS2 / CTSPC		**Yes**	**No**	
Physical IPV				
Perpetration — mild	Yes	13	2	
	No	1	100	0.95
Victimization —mild	Yes	14	4	
	No	0	99	0.93
Perpetration — severe	Yes	2	2	
	No	2	111	0.93
Victimization — severe	Yes	2	2	
	No	6	107	0.86
Psychological IPV victimization	Yes	21	6	
	No	0	4	0.61
Physical parent-child aggression —mild	Yes	59	3	
	No	9	26	0.75
Physical parent-child aggression — severe	Yes	4	3	
	No	0	91	0.94
Psychological parent-child aggression	Yes	29	51	
	No	0	16	-0.06

Note: CTS2 = Revised Conflict Tactics Scale; CTSPC = Parent-Child Conflict Tactics Scale; FM = Family Maltreatment measure.

[a]For breakouts of exact items endorsed when FM and CTS2/CTSPC were discordant see Supplemental Tables 2a–2h

Table 2.

Comparison of CS-IPV and CS-child abuse from FM and coded recounted incidents

	Verbalized incident			
Family Maltreatment Measure	**Met Criteria**	**Did Not Meet Criteria**	**G**	**No Audio- Recounted Incident Provided**
Physical CS-IPV perpetration				
Met Criteria	1	1		3
Did Not Meet Criteria	0	5	0.71	115
Physical CS-IPV victimization				
Met Criteria	2	0		2
Did Not Meet Criteria	0	11	1	110
Psychological CS-IPV victimization				
Met Criteria	2	5		14
Did Not Meet Criteria	0	0	−0.43	103
Physical CS-child abuse perpetration				
Met Criteria	1	18		15
Did Not Meet Criteria	0	58	0.53	156
Psychological CS-child abuse perpetration				
Met Criteria	0	10		5
Did Not Meet Criteria	0	31	0.51	202

Notes. For coding of verbalized incidents, “met criteria” = the participant described an act of maltreatment *and* described an impact that met criteria; “did not meet criteria” = the participant (a) described an act of maltreatment and (b) explicitly indicated that there was no impact; OR (a) described an act of maltreatment, (b) did not describe an impact, and (c) described an act that was highly unlikely to have met criteria for impact.