

Accepted by 'Aggression and Violent Behavior' on 29.08.2017

A Critique of the Revised Conflict Tactics Scales-2 (CTS-2)

RICHARD TOBY JONES¹

lwxtjo@nottingham.ac.uk

KEVIN BROWNE¹

Kevin.Browne@nottingham.ac.uk

SHIHNING CHOU¹

Shihning.Chou@nottingham.ac.uk

¹Centre for Forensic and Family Psychology, Division of Psychiatry and Applied Psychology, School of Medicine, University of Nottingham, Floor B, YANG Fujia Building, Jubilee Campus, Wollaton Road, Nottingham, NG8 1BB, UK.

CORRESPONDING AUTHOR:

SHIHNING CHOU

Centre for Forensic and Family Psychology, Division of Psychiatry and Applied Psychology, School of Medicine, University of Nottingham, Floor B, YANG Fujia Building, Jubilee Campus, Wollaton Road, Nottingham, NG8 1BB, UK.

Email: Shihning.Chou@nottingham.ac.uk

Telephone: +44 (0)115 84 66623

A Critique of the Revised Conflict Tactics Scales-2 (CTS-2)

ABSTRACT

The purpose of this review is to critically evaluate the Conflict Tactics Scales-2, a measure used worldwide for research and clinical practices. The reliability, validity and normative samples of the CTS-2 are considered and compared with similar psychometric measures. The limitations of the original CTS have been discussed alongside the CTS-2. Reliability is considered to be good to excellent by alpha coefficient and the variance explained by differing samples or methods of administration. Caution is recommended when interpreting the CTS-2 in clinical settings. Researchers point toward a five-factor structure to the CTS-2. The importance of appropriate norms is discussed and considered crucial when using the tool in settings where reporting patterns may differ. Difficulties in comparing CTS-2 scores across samples, cultures and countries are highlighted. Overall, the CTS-2 is a robust psychometric measure, although it holds limited clinical utility if it is used separately from other sources of information gathering (i.e. psychometric measures or interview). In order to enhance clinical utility, it should be administered alongside measures or clinical interviews that can provide added context regarding violence in the family. More research is required in diverse population samples, cultures/countries and languages.

KEYWORDS

CTS-2, CRITIQUE, FAMILY CONFLICT, VALIDITY, RELIABILITY

1. INTRODUCTION

1.1 Measuring Family Conflict

The 'dark figure' of violence within the family home is an issue for victimisation surveys and police recorded data (Sinha, February 2013). It leads to an underestimation of the extent of the problem and its cost to services. Therefore, measures of family conflict that enhance accuracy and encourage self-report are essential to those working in clinical and research settings. Although accuracy is desired, it is seldom achieved as there is often conflict of interest or social desirability bias associated with self-reported family conflict. However, through researchers offering complete confidentiality and anonymity, this can help encourage full disclosure (Hamby, 2014).

Measuring conflict in the family poses a multitude of problems to researchers and clinicians. The following can impact the usefulness of the data collected; a lack of disclosure from the involved couple, a lack of therapist investigation, confusion over terminologies used and only one partner being consulted (Straus, Hamby & Warren, 2003). Thus, tools are required to provide a universal and consistent approach to measuring family conflict. In particular, identifying the frequency and severity of family conflict can help determine the provision of appropriate services for the perpetrators and victims.

The Conflict Tactics Scales (CTS-2; Straus, Hamby, Boney-McCoy & Sugarman, 1996) is the most widely applied research and clinical tool in measuring family conflict. Other well-known tools include Partner Abuse Scale (PAS; Hudson, 1990), Composite Abuse Scale (CAS; Hegarty, Sheehan & Schonfeld, 1999), the Abusive Behaviour Inventory (ABI; Shepard and Campbell, 1992) and the Index of Spouse Abuse (ISA; Hudson & McIntosh, 1981).

1.2 Review Aims

Given the prevalence use of the CTS-2, this paper provides a critical review of the psychometric properties and practical utility of this tool. Furthermore, the paper attempts to evaluate whether the CTS-2 has improved upon the limitations of the CTS and compares the CTS-2 with similar measures.

1.3 The Original CTS

The Conflict Tactics Scales (CTS) were initially developed by Straus (1979) in response to a growing demand to quantify the extent of conflict in the family. This tool was intended to measure the use of reasoning/negotiation, verbal and physical aggression within the family household between partners in dating, cohabiting or marital relationships (Straus *et al.*, 1996). Straus, Hamby and Warren (2003) argue this focus on explicit actions yielded one of the many strengths of the CTS. The CTS were theoretically driven by conflict theory (Straus, 1979). This theory stipulates that human existence goes hand in hand with conflict, yet violence is not always an inevitable method adopted by an individual in order to quell conflict (Straus, Hamby & Warren, 2003). Straus (1979) envisioned the CTS to be one aspect of a battery of other psychometric measures which could assess more subtle facets of family violence i.e. fear or controlling behavior.

The CTS measure consisted of 19 items with three scales; (1) Reasoning (3 items), (2) Verbal Aggression (6 items) and, (3) Physical Assault (9 items) and one item which was not scored "cried". The measure was most frequently used to obtain information regarding physical assaults against both partners (Straus *et al.*, 1996). Therefore, the 19 items were asked twice giving a total of 38 items administered to each respondent. Its intended use was to measure the extent to which specific conflict tactics had been used in the family home (*ibid*). It was developed to be used by clinicians as an interview schedule as opposed to a self-administered questionnaire. The CTS schedule had two columns with the response categories (0-6), one for the number of times the respondent had carried out an act and one for the partner's acts over the past year. These response categories were coded 0 for never, '1' for once, '2' for

twice, '3' for 3-5 times, '4' for 6-10 times, '5' for 11-20 times, '6' for more than 20 times and 'X' for don't know. There was also a section to record whether the conflict tactic had ever happened (i.e. not in the past year). The item order began with items of low coerciveness (e.g. reasoning items) with gradually more coercive conflict tactics towards the end of the measure (Straus, 1979).

Since its creation, the CTS have been utilised worldwide in culturally diverse clinical and normative population samples (Giles-Sims, 1983; Tjaden & Thoennes, 2000; Straus *et al.*, 2003). The measure has been implemented within family therapy to help detect domestic violence (Aldorando & Straus, 1994; Boughner, Hayes, Bubenzer & West, 1994; Straus *et al.*, 2003; Sherman & Fredman, 2013). Some authors have also adapted the CTS in order to measure different aspects of family violence, demonstrating its flexibility and utility for research purposes, specifically criminal justice related research (Browne, 1987; Gondolf, 1988; Straus, 1993).

The CTS was considered innovative due to the measure's focus on behavioral descriptions of conflict (e.g. punched) and avoidance of the words 'abuse' and 'rape' which could be interpreted differently by respondents (Hamby, 2014). However, since its inception the CTS measure has become the centre of controversy related to the perpetration of partner violence. Research that has utilised the CTS has found support for partner violence being perpetrated by both genders to significant levels (Archer, 2000; Straus, 2010), a pattern termed 'gender symmetry' (Winstok, 2015). Such findings contradicted the data from other types of surveys (e.g. victim surveys, police recorded data and national surveys) that show gender 'asymmetry' where men predominantly perpetrate partner violence in their intimate relationships (Winstok, 2015).

This controversy has spawned many 'first generation' critiques of CTS and symmetrical gender patterns in partner violence (e.g. Dobash & Dobash, 1979; Tolman, 1989; Saunders,

1992; Campbell, 1995; cited in Hamby, 2015). There have also been a number of current critiques on gender patterns in partner violence and we are no closer to a definitive answer (Kimmel, 2002; Hamby, 2014, Winstok, 2015, Hamby 2015, Winstok & Straus 2016).

Dobash and Dobash (2004) highlighted concerns about the CTS' external validity with the meaning of certain behavioral acts being open to interpretation (e.g. thrown an object at your partner) and the context behind the act (e.g. retaliation or self-defence). Additionally, the same authors argued that if a partner committed one act they are equally defined as 'violent' as the other partner who may have committed several acts. Overall, the authors argued that the measure,

"...distorts the reality of intimate partner violence..." (Dobash & Dobash, 2004; pp. 330).

Further definitional problems were noted by Dobash and Dobash (2004) such as the measure defining acts on the psychological/emotional abuse subscale 'abusive' without knowing the context in which the behavior occurs (e.g. threatening to leave). Another problem of the CTS is it's merging of violent acts (e.g. physical and sexual acts) with non-violent acts of abuse (e.g. name calling and shouting) and referring to it interchangeably as either violence or abuse (Dobash & Dobash, 2004). This can lead to misleading findings and confusion over what actually constitutes as violent acts. Dobash and Dobash (2004) argued these violent and abusive acts should be appropriately separated, studied and described. Overall, Dobash and Dobash (2004) concluded the CTS' act based approach to measuring partner violence was too narrow and more likely to find gender symmetry because it does not include the context, consequences, motivations and intentions behind the partner violence.

Further limitations of the CTS had also been noted by critics. Firstly, the measure had received criticism regarding the number of violent acts that were included in forms (Straus, 1987). Critics have argued that this restriction misrepresents the extent of violence in the family home as they are a multitude of conflict tactics not accounted for. Secondly, expecting an accurate

recall from respondents from the past year had been highlighted as a problem, if conflict tactics experienced, are common. Calls for a shorter recall period have been made (Straus, 1987). Thirdly, the measure did not distinguish the severity of violent acts in the scoring and interpretation (e.g. a slap was considered in the same category of seriousness as using a knife). Fourthly, the measure did not account for injury or sexual coercion. Fifthly, Straus (1987) criticised the measure for having an inadequate number of items to measure verbal aggression and reasoning tactics. Finally, critics have emphasised that the measure was created on the assumption that violence is family-based and not male orientated (DeKeseredy & Schwartz, 1998).

1.4 CTS-2 Development

The CTS-2 was developed by Straus et al. (1996) to address some of the limitations of the CTS. In particular, a greater distinction between minor and severe conflict tactics, increasing the number of acts measured (incl. verbal aggression and reasoning/negotiation), measuring injury and sexual coercion, and clarifying terms used in the measure were improvements that were targeted by the authors tasked with revising the CTS. The revision of the CTS involved reformatting the design of the measure to improve the presentation and intersperse items to minimise demand characteristics (Straus *et al.*, 2003). The CTS was initially developed as an interview schedule rather than a self-administered measurement tool, therefore, the presentation was simplified in the CTS-2 (Straus *et al.*, 2003). Therefore, the criticisms the original measure received regarding the clarity of the presentation may have been unfair. Straus et al. (1996) focussed on item wording in the revision of the CTS and the ambiguity of items such as ‘threw something at him/her’ and changing it to ‘threw something at my partner that could hurt’. Developing clarity of item wording was thought to help to prevent false-positive results as consensual acts such as pillow fights could fit with the original item wording and they are clearly not related to rates of partner violence (Hamby, 2015). Further, the item wording was also revised to provide gender neutral terms such as ‘my partner’ so that the measure was applicable for use to explore conflict tactics used in same-sex relationships

(Hamby, 2015). Items were also removed or combined to improve distinction between each item (Straus *et al.*, 1996). This process was guided by the critiques of the CTS, clinical experience of the authors and existing research (Straus *et al.*, 2003). The CTS-2 authors used a specific set of criteria to determine the selection items for the measure. For example, specific behaviors such as animal cruelty or child abuse were excluded as these behaviors could only occur if the couple have children or pets (*ibid*). Overall, the items were designed to be inclusive, specific to acts or events, cover a range of severities and be understandable in terms of reading ability (*ibid*).

In the revisions by Straus et al. (1996). The Reasoning scale was renamed Negotiation as the authors believed this term better described the construct measured by the items (Straus *et al.*, 2003). Further, the Verbal Aggression scale was changed to Psychological Aggression as some of the items related to non-verbal acts of aggression (*ibid*). The items in these original scales were added to in the revised version of the CTS. This was to increase the scope and to respond to criticisms relating to the brevity of the measure (Straus, *et al.*, 2003).

The CTS-2 incorporated two new scales, Injury and Sexual Coercion. The CTS-2 authors chose these additional scales due to their fundamental importance in relation to partner violence (Straus *et al.*, 1996). Sexual coercion is another form of partner abuse that was not covered by the original CTS and therefore, added to the scope of the measurement tool. Although, critics would argue that the scope of sexual coercion could have been increased if items specified more non-penetrative acts (Hamby, 2014). The injury scale was also added to increase the scope of the measure as it highlighted the consequences on the partner (Straus *et al.*, 2003). Both scales fitted well into the framework and design of the CTS-2 as they focussed on 'act based' behavior and had similar response categories to the other scales (e.g. how often did this happen in the past year?). However, it could be argued that these scales were included in response to 'gender symmetry' criticisms made regarding the original measure (Hamby, 2015). Since, its development these new scales have provided evidence of

'gender asymmetry' with higher levels of sexual coercion and more severe injuries against female partners (*ibid*). Items in each scale were grouped due to content covered in those items (Straus *et al.*, 2003). For example, negotiation items covered emotional and cognitive tactics and the other scales of items distinguished between minor and severe levels. Again, it can be argued that these revisions were thought to shift gender patterns found by the original CTS (Hamby, 2015).

Overall, it was hoped that the newer version would extend the scope of the tool in clinical settings (Straus *et al.*, 2003). Each version of the measure required the item to be answered once by the respondent and once regarding the respondent's perception of their partner's behavior. Therefore, the number of items is doubled. The original CTS had a total of 38 (19) items and the revised CTS-2 had 78 (39) items.

Despite its revision, the CTS-2 measure had attracted further criticism, some of which relate to the original measure. Hamby (2014) who was tasked with the redevelopment of the original CTS, highlighted that the measure in one form or another had been used for over forty years with limited changes made to the instrument. The author uses this point to argue that the 'technology' has now become outdated and researchers must develop new technologies to improve the measurement of violence (Hamby, 2014). However, this argument has been countered on the notion at the age of the technology should not be the determining factor on whether it is outdated (Winstok, 2015). Additionally, Winstok (2015) argued that a valid and reliable alternative to the CTS-2 measure had not been presented by Hamby (2014).

Hamby (2014) uses conservatism in academia as an argument against the CTS-2 measure. Specifically, Hamby (2014) argued that journal and grant reviewers have preferred research articles that confirm their own pre-conceptions about partner violence and are more likely to accept research that adopts frequently used measures (e.g. CTS-2). Similarly, Hamby (2014) suggested that some authors look to publish many articles from a single project and are

incentivised to publish quantity over quality. According to Hamby (2014), these issues have created a 'systematic drag' on science and thus our knowledge on partner violence. In contrast, Winstok (2015) argued that it is those who have critiqued the CTS have been favoured by journal reviewers and that authors studying violence against women have received more funding. These issues may reflect the biases amongst most research areas and not just that of partner violence and the CTS-2 measure. Nevertheless, it is important to recognise these wider perspectives to remain critical of the literature.

1.5 CTS-2 Overview

The CTS-2 is a 78 item measure which assesses the amount of conflict between partners in a dating, co-habiting or martial relationship (Straus *et al.*, 1996). Since its revision the CTS-2 has been one of the most widely used psychometric measures of Intimate Partner Violence (Yun, 2011). The CTS-2 can be adapted to suit different time frames, situations, relationships and response categories (Straus *et al.*, 1996). The scales on this measure are; Negotiation (six items), Psychological Aggression (eight items), Physical Assault (twelve items) – which were already incorporated into the CTS – with the addition of two new scales Injury (seven items) and Sexual Coercion (six items). Subscales were introduced for the Negotiation scale to differentiate between cognitive and emotional forms of reasoning tactics. Furthermore, the other scales provide minor and severe subscales.

The CTS-2 is scored by adding the midpoints of the response categories (e.g. for the 3-5 category this would be 4). The measure requests respondents to approximate the number of times an event occurred in the last twelve months either once, twice, 3-5 times (scored as 4), 6-10 times (scored as 8), 11-20 times (scored as 15) or more than 20 times (scored as 25). Further, there is an option to endorse whether the conflict tactic had ever occurred (never or not in the past year). This option is scored, tallied and interpreted separately (e.g. 0 if it has never occurred or 1 if it has occurred). Responses are then transferred onto the scoring worksheet with odd numbered responses for 'self' and even numbered responses for 'my

partner' (self and partner are scored separately). The researcher or clinician then tallies the item scores according to the scale the item belongs to (e.g. psychological aggression). Once the raw scores (past year) for each scale and partner are tallied they are transferred to the CTS-2 summary sheet. Total possible scores for each scale (past year items) are as follows; Negotiation 150, Psychological Aggression 200, Physical Assault 300, Injury 150 and, Sexual Coercion 175. Each of the scales raw scores are considered standalone and are not added together to provide a total score. These raw scores are considered an estimate of frequency of conflict tactics in the respondents' relationship in the past year.

Straus et al. (2003) emphasize that for clinical purposes, the measure should be used as behaviour checklist where any non-zero response on negative conflict tactics would result in follow-up inquiries to determine problem areas. However, Straus et al. (2003) also note caution as response will need to be verified following a robust clinical assessment using a range of sources (e.g. medical records, clinical interview etc.).

In Interpreting the CTS-2 scores clinically, Straus et al. (2003) suggest that for raw scores on the physical assault, injury and sexual coercion scales would warrant further investigation and appropriate steps would need to be taken to safeguard the victim. However, for the Negotiation scale the authors suggest that higher scores indicate a greater use of positive conflict tactics by the respondent. Conversely, lower scores indicate an absence of these alternatives to aggressive tactics which would mean respondents would need to develop and practice these skills. Regarding the Psychological Aggression scale, Straus et al. (2003) suggest that respondents that endorse higher frequencies may require therapeutic work to increase their awareness of the impact psychological aggression can have on their relationship.

However, for research purposes the CTS-2 tends to be interpreted differently. Straus et al. (2003) report that researchers would be most interested in using the CTS-2 to estimate prevalence and chronicity rates in specific samples. The prevalence rate would refer to the

percentage of the sample that reports at least one instance of each behaviour occurring (Straus *et al.*, 2003). Chronicity refers to the average amount the behaviour has occurred within the sample who endorse that they have used a particular conflict tactic type (e.g. psychological aggression) or specific behaviour (e.g. hitting) (*ibid*).

In comparison to the published research papers on the CTS, the CTS-2 has considerably less journal articles regarding its applicability and psychometric properties. Moreover, it has been nineteen years since the CTS-2 was developed. However, in this time, twelve papers have been published evaluating or critiquing aspects of the psychometric measure. These reviews focus on the following issues; factor structure (Connelly, Newton & Aarons, 2005; Calvete *et al.*, 2007; Yun, 2011), validity (Jones *et al.*, 2002; Straus, 2004; Relva, Fernandes & Costa, 2013) and reliability (Lucente *et al.*, 2001; Newton, Donaldson, Connelly & Landsverk, 2001; Dietz & Jasinski, 2007; Ro & Lawrence, 2007; Vega & O'Leary, 2007) in a variety of settings. Each paper has tended focus on one particular issue rather than taking an eclectic approach. An aim of this review is to build upon the previous researchers views on the CTS-2. The authors of CTS-2 have also published the psychometric properties of the measure in research articles (Straus *et al.*, 1996; Straus, 2004).

For research purposes the CTS-2 provides an estimation of Partner Violence (PV). However, if this measure was used in criminal justice proceedings it should not be the only psychometric administered as it has no validity scale built into the measure. Straus (1993) noted that researchers need to examine the CTS-2 under different population samples given its potential to provide changeable psychometric properties. Furthermore, amongst the research literature the CTS has been used controversially to investigate asymmetry or symmetry of prevalence amongst male and female perpetrators. This outlines one of the potential misuses of the tool (Dobash *et al.*, 1992; Dietz & Jasinski, 2007).

2. RELIABILITY

The CTS-2 poses a number of caveats to estimating reliability. Firstly, the assessment does not measure stable characteristics like a relatively stable personality construct, instead, it measures behavior which can change over time (Vega & O'Leary, 2007). Secondly, by assessing actual behavior, the level of occurrence may not be normally distributed (*ibid*). Therefore, creating a statistical problem, as the base rates for the behavior could be either high or low (*ibid*). Finally, differences in inter-rater correlations may be an inaccurate portrayal of reliability (*ibid*) for the reasons set out below.

The inter-rater reliability of the CTS-2 has been found to range between low and medium (O'Leary & Williams, 2006). This issue is highlighted within court based assessments, although it would be recommended that both partners should complete the psychometric to assess consistency of reporting. Socially desirable responses or a desire to avoid negative consequences are more probable within this clinical setting.

Few studies have measured the CTS-2 on the grounds of test-retest reliability. Researchers focus has largely been on the factor structure, internal consistency and the validity of the measure (Vega & O'Leary, 2007). In Vega and O'Leary's (2007) study they measured incarcerated males' responses on the CTS-2 on two occasions regarding the same time period analysing the stability of self-reported aggression. The authors reported strong test-retest reliability across four of the aggression subscales for both the participants own behavior and their partner's (Vega & O'Leary, 2007). The test retest period was over two months which is relatively long compared to similar studies testing test-retest reliability (Birchler & Fals-Stewart, 1994; Mills, Kroner & Forth, 1998). However, this period may still have not taken into account the propensity for someone to report less aggression tactics over longer time frames, particularly as the participants in Vega and O'Leary's (2007) study showed signs of underreporting at both test periods. Furthermore, the authors of the paper did not analyse the other partner's responses.

Lucente, Fals-Stewart, Richards and Goscha (2001) measured the internal reliability of the CTS-2 using the alpha coefficients for the individual subscales and the test overall on a population of incarcerated females. Each subscale was measured to have high levels of consistency ($>.74$) apart from the perpetration sexual coercion subscale (*ibid*). In contrast, Newton et al. (2001) found a low reliability score for the physical assault scale. Overall, these results provided a strong case for the internal consistency of the CTS-2 scales. However, the authors suggested further research could analyse the minor and severe items of the CTS-2 (Lucente *et al.*, 2001).

Yun (2011) analysed the minor and severe items, noting adequate internal consistency of .68-.85 for perpetration and .68-.84 for victimization. However, Yun (2011) found the minor sexual coercion alpha coefficient to vary from .18 (perpetration) - .37 (victimisation). Furthermore, due to its clinical use, Yun (2011) argued that the CTS-2 alpha coefficients should be cautiously accepted. Otherwise, influential decisions could be based upon reliability scores which have not reached the desired standards of .9. The overall quality of Yun's (2011) findings are limited due to a self-selection bias and a sampling pool from south-eastern US states only. Furthermore, the following subscales; Physical Assault, Sexual Coercion and Injury subscales were heavily skewed and kurtotic (Yun, 2011).

Like all the studies mentioned above, each used data from one partner and did not compare these with data from their partners. Therefore, there were potential response biases. Additionally, the majority of these papers had collected data on specific populations, such as female prisoners and high risk postnatal females, which lack generalisability across other samples (Dietz & Jasinski, 2007). Despite this, alpha coefficients have been found to range between .79 and .95 (Straus *et al.*, 1996). In a larger cross-cultural study conducted in 33 countries with 6,000 students these coefficients ranged between .74 and .89 for each subscale (Straus, 2004). These coefficients demonstrate good to excellent internal reliability.

In comparison to the original CT scales, Straus et al. (1996) reports similar or higher internal consistency in the CTS-2. This suggests that along with making aesthetical improvements the changes made by the authors have not effected its good reliability. Comparatively speaking, other measures such as the CAS have demonstrated similar levels of internal reliability, yet more consistently over each dimension >0.85 (severe combined abuse, emotional abuse, physical abuse and harassment) (Hegarty, Bush, Sheehan, 2005). Moreover, the ISA has been found to have alpha coefficients over .90 within a sample of African-Americans (Campbell, Campbell, King, Parker & Ryan, 1994).

Ro and Lawrence (2007) compared the psychological aggression subscale of the CTS-2 to two other measures of psychological aggression, the Multidimensional Measure of Emotional Abuse (MMEA; Murphy & Hoover, 1999) and the Test of Negative Social Exchange (TENSE; Ruehlman & Karoly, 1991). The authors found that despite its wide use in the literature, the CTS-2 subscale had lower internal consistency than the other measures. Ro and Lawrence (2007) opined that the MMEA included a wider array of behaviors than the CTS-2 subscale and considered that the CTS-2 subscale should not be used in studies where psychological aggression would be the sole focus.

Reichenheim, Klien and Moraes (2007) argued that the CTS-2 evaluative literature has focussed on the classic test theory item analysis compared to the robust item response theory approach. An item response theory approach focusses on the item by evaluating the questionnaire responses rather than the whole test itself. It is considered by researchers to be a more in-depth analysis (Reichenheim *et al.*, 2007). Further, the authors noted that no researchers had used this approach with CTS-2 data regarding heterosexual couples. Reichenheim et al. (2007) focussed on the Physical Assault scale and found support for a single factor solution. This confirmed the measure's good to excellent internal consistency and highlights that the items discriminated well from each other. However, the authors also found

that items on the scale may not indicate the same levels of violence for males and females. This discrepancy was noted to

“...detract from the scale’s universality” (Reichenheim et al., 2007; pp 58.).

Researchers have investigated the influence item order may have on IPV rates (Ramirez & Straus, 2006; Dietz & Jasinski, 2007). Ramirez and Straus (2006) indicated that modifying the item order can elicit greater disclosure rates for males and females, yet switching the pair order of questions had no effect. However, the order of the items was not randomised and they used a small sample size (Dietz & Jasinski, 2007). Logistic regression analyses did not indicate any significant concerns with the four versions of the CTS-2 examined in Dietz and Jasinski’s (2007) study. Therefore, the question format may not be critical if the researcher’s aim is to measure physical aggression and IPV. Conversely, if the researcher’s goal is to measure the concepts of psychological aggression or sexual aggression they may wish to place the negotiation scaled items first as this item order seemed to enhance the reporting of these concepts (*ibid*). Regardless, the item order differences may be minor statistically, though clinically these differences are meaningful especially when some concepts relate to lower base rates such as sexual coercion (*ibid*). This research lacked generalisability to other populations based on the use of an opportunistic sample of students. Despite this, it allowed for a larger sample size from which multiple comparisons could have been made. Overall, the researchers highlight that the unpaired (where the respondent answers based on his/her behavior and then their partner’s on the following item) version of revised CTS-2 is able to more accurately determine the prevalence of IPV (*ibid*).

3. VALIDITY

Jones, Ji, Beck and Beck (2002) found an inverse correlation between the CTS-2 negotiation subscale and the Abusive Behavior Checklist (ABC; Beck & Beck, 1998, cited in, Jones et al., 2002). Somewhat unexpected, the authors found that the more negotiation tactics a victim used was associated with increased abusive behavior towards them from their partner (*ibid*).

This suggests the measure can compare and correlate with similar measures of partner violence to provide additional understanding or context. In this case, incarcerated females coped with their abusive environments by employing more negotiation tactics (Jones *et al.*, 2002). However, the ABC was found not to significantly correlate with the Injury scale. Jones *et al.* (2002) argued that the reason behind this was that the ABC is a measure of physical assault behaviors whereas the Injury subscale measures the effects of the physical assault. Conversely, the General Assault construct (consisting of the Psychological Aggression and Physical Assault) positively correlated with the ABC suggesting concurrent validity (*ibid*). A methodological limitation in this study related to the researchers asking the incarcerated females to recall conflict tactics from an average of four years prior (*ibid*). This calls into question the validity of their results and whether it is appropriate to test offenders who have been serving a prison sentence for longer than a year. Furthermore, Straus and Douglas (2004) compared the CTS-2 to a short form version of the CTS-2. In their study they were able to determine the concurrent validity of the measures. For perpetration the correlations ranged from .77 to .89 and for victimization .65 to .94 indicating strong positive relationships between the measures. However, due to the short form using the highest correlated items from the CTS-2 it is likely that these concurrent validity coefficients are inflated.

Researchers employing different versions of the CTS such as the sibling version have found evidence for construct validity. The CTS-2 was found to positively correlate with the sibling and parent-child versions (Relva, Fernandes & Costa, 2013). However, the sample was not representative of the Portuguese general population and retrospective reporting was used (Relva *et al.*, 2013). Tiwari, Fong, Chan, Leung, Parker and Ho (2007) found only fair kappa coefficients between the Abusive Assessment Screen and the Chinese version of the CTS-2 suggesting discrepancies in participant responding across similar measures. However, this could be attributed to cultural differences.

There are limitations to the CTS-2 relating to content validity. These are issues that were identified as areas that the CT scales could have improved upon. As a construct which measures levels of partner violence, the CTS-2 does not include the motives of the violence, this would inevitably add context to the calculation of frequency. Furthermore, definitions of partner violence often involve the use of controlling behavior or coercive tactics (i.e. financial power/isolation). This is another avenue not covered by the current CTS-2 or improved on from the original CT scales. These are areas which other measures may be able to explore when used with the CTS-2. Therefore, the CTS-2 should not be recommended as the solitary questionnaire to explore partners' conflict tactics. If so, there would be a danger of losing key information relating to the intensity of the partner violence or more covert aspects of domestic violence tactics.

A five factor model underpinning the CTS-2 has been established through confirmatory factor analyses (Newton, Connelly & Landsverk, 2001; Straus, 2004; Connelly, Newton & Aarons, 2005; Calvete, Corral & Estevez, 2007). The goodness of fit index used by Newton et al. (2001) indicated that the five factor model was an acceptable fit ($>.9$). Furthermore, Calvete et al.'s (2007) goodness of fit indices suggested a good model. This indicated that Straus et al.'s (1996) understanding of the nature of conflict tactics is consistent with the measures of the construct. Additionally, the distinctions between minor and severe forms of violence were supported by confirmatory factor analyses (Calvete *et al.*, 2007). This corresponds with previous research (Hamby & Sugarman, 1999; Newton *et al.*, 2001). Furthermore, the original CT scales was found to support a two factor model (psychological and physical aggression), the CTS-2 demonstrates an improvement on this, due to the increased scope of the measure (Cascardi, Avery-Leaf, O'Leary & Slep, 1999).

However, Calvete et al. (2007) were the only researchers to focus on the victimisation form and not the perpetration form. Further, the authors included all scales in the analyses. Connelly, Newton and Aarons (2005) examined the factorial validity of the CTS-2 using

different languages. They noted caution when comparing results of English and Spanish language groups because of the potential nuances in patterns of responding owing to cultural, educational and economical differences (*ibid*). These studies all used female samples, therefore more research is needed to confirm the factor structure of the CTS-2 using male samples.

The studies investigating the factor structure of the CTS-2 were all administered to specific populations and only included some of the total items (Yun, 2011). Yun (2011) investigated the factor structure using a ten factor model because of the minor and severe item groupings on the CTS-2 scales and the sub-constructs on the Negotiation construct (Straus *et al.*, 1996). Yun (2011) demonstrated that violence scales were all significantly correlated with one another. However, the factor loadings did not load exclusively on minor or severe item groups. The complex factor loadings found support for an eight factor model instead as the Negotiation sub-constructs were found to be independent (*ibid*). Overall, this suggests that the CTS-2 factor structure does not support the minor and severe item groupings hypothesized by Straus *et al.* (1996), though Yun (2011) suggested that additional confirmatory analyses are needed to investigate this further.

4. NORMS

It is crucial for a psychometric measure to provide a wide range of normative samples, particularly if the measure is used in a clinical context. For example, incarcerated women were found to report higher rates of chronicity and severity of abuse compared to Straus *et al.*'s (1996) sample of college women. In a sample of Latina adolescents, Newman and Campbell (2011) found that the sample had a tendency to match or exceed the level and frequency of the violent behaviors against their partners.

Archer (1999) found in a meta-analysis of 43 studies that participants underreported aggression on the CTS regardless of gender, demonstrating the issue of social desirability. This is an area where more research is required so that the tools can be compared effectively. As it stands, issues with respondent disclosure remains a limitation with the CTS-2. One of Straus et al's (1996) aims was to improve spousal agreement of the CTS with the CTS-2. However, this has not been found in Simpson and Christensen's (2005) research where partner agreement was noted to be between poor to moderate for individual violent tactics and scales. The authors considered that this was either a limitation of the CTS-2 scale or an inevitable problem faced by those measuring partner violence (*ibid*). Either way, the revised CTS-2 does not seem to have improved from its predecessor in this aspect.

Hamby (2014) argued that research using the CTS-2 had used non-representative samples such as small convenience samples of college students. Further, Hamby (2014) noted that those taking the gender symmetry perspective favoured research with small samples as opposed to nationally representative and population based surveys that had highlighted asymmetric gender patterns (e.g. Archer, 2000). However, gender symmetry in perpetration was also noted in national and international samples countering this argument (Gelles & Straus, 1988; Straus *et al.*, 2006; Straus, 2008; cited in, Winstok, 2015).

The CTS-2 has been utilised in a variety of countries with the form having been translated into different languages (see Yan & Tang, 2001; Straus, 1999). This may pose a number of issues relating to external validity as the psychometric properties of the CTS-2 might not be comparable in other countries especially where there are cultural and ethnic differences. Although there are a number of published studies which focus on the psychometric properties of the English version of the CTS-2, there is a lack of published studies investigating the psychometric properties of the CTS-2 in non-English speaking populations (Connelly, Newton & Aarons, 2005). This outlines the importance of investigating the scientific properties of the CTS-2 scales in ethnically diverse and non-English speaking populations, especially, in this

case, with a widely used measure such as the CTS-2. Straus (2004) found the CTS-2 could be used cross culturally to measure violence in dating relationships in university students. However, it was acknowledged that this could not be extended to other applications of the measure or for use amongst different sample populations as of yet (*ibid*). Connelly et al. (2005) were one of the first to do this and have called for additional research particularly as the internal consistency and factor structure model were weaker in the Spanish CTS-2. Additionally, the authors noted that most of the published research focussing on the factor structure of the CTS-2 were conducted with female samples (*ibid*). Nevertheless, Connelly et al. (2005) argued in support of using the CTS-2 in countries with differing languages and diverse populations.

The Spanish version of the tool was found to be valid amongst a sample of women who completed the victimization form (Calvete *et al.*, 2007). The authors noted that there were significant differences in the prevalence rates, chronicity and level of frequency for all scales except the minor psychological aggression and sexual coercion scales when comparing the results from different samples. This highlights the importance of obtaining appropriate norms and selecting the 'best fit' sample to match the participant's age, setting and other demographic variables. Relating to discriminant validity, the authors were able to identify a subset in their sample of 1296 Spanish women who were accessing victim service compared to the rest of the sample as they were reporting higher rates of victimisation (Calvete *et al.*, 2007; Newman & Campbell, 2011). Wider research is needed in non-westernised countries to draw comparisons with and to investigate the CTS-2's validity further.

5. ENCODING

Initially, the handbook and CTS-2 summary sheets provided an average CTS-2 scores (chronicity) and prevalence rates by gender and whether or not the respondent was reporting their own perpetration or victimization. These norms were based on a college student sample

of 113 men and 204 women (Straus *et al.*, 2003). If these figures are to be judged as typical, the prevalence rates are likely to be in the range of: 98-100% for Negotiation; 74-83% for Psychological Aggression; 31-49% for Physical Assault; 9-16% for Injury and; 18-38% for Sexual Coercion. For chronicity, the range of averages reported per year were: 57.4 – 69.7 for Negotiation; 15.1 – 17.2 for Psychological Aggression; 9.3 – 15.9 for Physical Assault; 3.6 – 25.1 for Injury and; 11.8 – 19.9 for Sexual Coercion.

Since its development, the CTS-2 has been used to determine prevalence rates of partner violence in a number of countries and samples. Langhinrichsen-Rohling, Selwyn and Rohling (2012) conducted a review of partner violence prevalence rates across samples, sexual orientations and ethnicities. 54% of the included studies used the CTS-2 or a version of the CTS-2 to assess violence rates. Langhinrichsen *et al.* (2012) found the samples identified significantly different prevalence rates of violence. Large population based samples indicated the lowest rates of violence with 22%. However, these studies used less questions derived from the CTS-2. In comparison, the following samples demonstrated higher partner violence prevalence rates; gay lesbian and bisexual samples 32%, school samples 33%, community 42%, predominantly-female treatment-seeking samples 77%, and approximately 100% of predominantly-male legal/justice samples (*ibid*). These findings were comparable with Straus' (2008) large international student sample from 32 nations who completed the CTS-2 with a prevalence rate of 31.2%. Furthermore, both Langhinrichsen-Rohling *et al.* (2012) and Straus' (2008) findings appear consistent across time with a prevalence study from 34 years ago indicating partner violence rates of 30% (Bernard & Bernard, 1983; cited in Langhinrichsen-Rohling *et al.*, 2012). Furthermore, the CTS-2 measure has identified the prevalence of bidirectional and unidirectional violence. Across the samples identified in Langhinrichsen-Rohling *et al.* (2012), the average amount of bidirectional violence was 57.5%. This indicates that partner violence is predominantly bidirectional (i.e. both partners have used partner violence to a certain degree within the relationship).

In addition to this research on prevalence rates, Straus (2015) used the CTS-2 to classify family dyads into the following categories; (1) both aggressed partner violence; (2) one partner aggressed only and; (3) the other partner aggressed only. Researchers refer to these categories as Dyadic Concordance Types (DCTs). In relation to heterosexual intimate relationships the dyads are; *Male-Only*, *Female-Only*, *Both-Aggressed* and *Neither-Aggressed* categories. Straus (2015) reported the DCTs prevalence rates using Miller et al's (2011) World Mental Health Survey data from eleven nations. This highlighted the following prevalence rates; male-only 36-39%, the female-only 19-21% and, both 42-43% of the 3,642 couples that reported physical assaults had occurred. The above figures were in relation to physical assaults only and not the other forms of partner violence measured by the CTS-2 such as sexual coercion. However, Straus' (2008) study of international dating violence in 32 nations highlighted that the male-only DCT was not the most prevalent category (even in the minor and severe violence categories). For the prevalence of psychological aggression and sexual coercion, similar findings have been identified (Straus & Michel-Smith, 2014; Straus & Michel-Smith, 2015). Overall, the predominant pattern found was that both partners were violent.

6. CONCLUSION

Overall, the CTS-2 has an extensive literature outlining its psychometric properties. However, this literature is not as extensive as the previous version of the tool (Straus *et al.*, 1996). The CTS-2 has good scientific properties, though this is largely because of research conducted in the US. Generally, its alpha coefficient on all subscales range between good to excellent, suggesting that the tool is internally valid. Furthermore, there is strong support for a five factor model for the CTS-2 suggesting good construct validity supporting Straus et al.'s (1996) theoretical hypotheses. Elements of the CTS-2 have been found to correlate well with other measures used for exploring abusive behaviors and psychological aggression indicating good convergent validity (Jones *et al.*, 2001; Ro & Lawrence, 2007). No other measure of conflict

tactics incorporating an array of conflict tactics have been developed with the similar levels of research papers. The CTS-2 is a robust measure for measuring frequency and chronicity of partner violence despite clear difficulties regarding patterns of responding (i.e. social desirability and partner agreement).

However, the CTS-2 is not without its criticisms. The focus of CTS and CTS-2 research has been largely on proving or disproving gender symmetry of partner violence, specifically, physical assault. This has had unfortunate consequences on our knowledge of the CTS-2 measure. For example, this has stunted investigation into the following scales; negotiation, psychological aggression, sexual coercion and injury. Research involving more specific populations provide more variance regarding the internal consistency of the CTS-2 scales (Yun, 2011). This suggests an unacceptable level of reliability if the measure is to be used to make important decisions in a clinical setting. Furthermore, other measures such as the ISA and CAS have been found to be at least similar, if not higher and more consistent in reliability, than the CTS-2 (Campbell *et al.*, 1994; Hegarty, Bush, Sheehan, 2005). In particular, the content validity could be improved upon otherwise the tool lacks context when exploring IPV. If the tool is used for clinical purposes other psychometric measures or clinical interviews should be considered to provide information on the motivations behind abusive behaviors or investigate the more covert conflict tactics such as financial control. This is a limitation that has not been improved upon from the original CT scales. Additionally, confirmatory factor analyses should be conducted to explore a ten factor model proposed by Straus *et al.* (1996). Further research is required to enhance the external validity and applicability of the CTS-2 across diverse populations, cultures, languages and countries. This is an area which needs addressing especially when the CTS-2 scales are used in other countries and translated in many languages. Therefore, additional clinical sample populations are required for more normative comparisons to be made.

ACKNOWLEDGEMENTS

I wish to thank Dr. Shihning Chou and Prof. Kevin Browne for providing their expertise and guidance on writing this article. I also wish to thank Sara Pugh for proof reading this paper.

REFERENCES

- Aldarondo, E., & Straus, M.A. (1994). Screening for physical violence in couple therapy: Methodological, practical, and ethical considerations. *Family Process, 33*, 425-439.
- Archer, J. (1999). Assessment of the Reliability of the Conflict Tactics Scales A Meta-Analytic Review. *Journal of Interpersonal Violence, 14*, 12, 1263-1289.
- Archer, J. (2000). Sex differences in physical aggression to partners: A reply to Frieze (2000), O'Leary (2000), and White, Smith, Koss, and Figueredo (2000). *Psychological Bulletin, 126*, 5, 697-702.
- Beck, M., & Beck, N. (1998). The Abusive Behavior Checklist. Unpublished instrument.
- Birchler, G.R., & Fals-Stewart, W. (1994). The Response to Conflict Scale: Psychometric properties. *Assessment, 1*, 4, 335-344.
- Boughner, S.R., Hayes, S.F., Bubbenzer, D.L., & West, J.D. (1994). Use of standardized assessment instruments by marital and family therapists: A survey. *Journal of Marital and Family Therapy, 20*, 1, 69-75.
- Browne, A. (1987). *When battered women kill*. New York: Free Press.
- Campbell, D.W., Campbell, J., King, C., Parker, B., & Ryan, J. (1994). The reliability and factor structure of the Index of Spouse Abuse with African-American women. *Violence and Victims, 9*, 3, 259-274.
- Cascardi, M., Avery-Leaf, S., O'Leary, K.D., & Slep, A.M.S. (1999). Factor structure and convergent validity of the Conflict Tactics Scale in high school students. *Psychological Assessment, 11*, 4, 546-555.
- Connelly, C.D., Newton, R.R., & Aarons, G.A. (2005). A psychometric examination of English and Spanish versions of the Revised Conflict Tactics Scales. *Journal of Interpersonal Violence, 20*, 12, 1560-1579.
- DeKeseredy, W. S., & Schwartz, M. D. (1998). Measuring the extent of woman abuse in intimate heterosexual relationships: A critique of the Conflict Tactics Scales. *US Department of Justice Violence Against Women Grants Office Electronic Resources*.
- Dietz, T.L., & Jasinski, J.L. (2007). The effect of item order on partner violence reporting: An examination of four versions of the revised Conflict Tactics Scales. *Social Science Research, 36*, 1, 353-373.
- Dobash, R.P., & Dobash, R.E. (2004). Women's violence to men in intimate relationships working on a puzzle. *The British Journal of Criminology, 324-349*.

- Dobash, R.P., Dobash, R.E., Wilson, M., & Daly, M. (1992). The myth of sexual symmetry in marital violence. *Social problems, 39, 1, 71-91.*
- Gelles, R.J., & Straus, M.A. (1988). *Intimate violence: The causes and consequences of abuse in the American family.* New York: Simon & Schuster.
- Giles-Sim, J. (1983). *Wife battering: A systems theory approach.* New York: Guildford Press.
- Gondolf, E.W. (1988). Who are those guys? Toward a behavioral typology of batterers. *Violence and Victims, 3, 187-203.*
- Hamby, S. (2014). Intimate partner and sexual violence research scientific progress, scientific challenges, and gender. *Trauma, Violence, & Abuse, 15, 3, 149-158.*
- Hamby, S. (2015). A Scientific Answer to a Scientific Question: The Gender Debate on Intimate Partner Violence. *Trauma, Violence, & Abuse, DOI: 1524838015596963.*
- Hamby, S.L., & Sugarman, D.B. (1999). Acts of psychological aggression against a partner and their relation to physical assault and gender. *Journal of Marriage and the Family, 61, 959-970.*
- Hegarty, K., Bush, R., & Sheehan, M. (2005). The composite abuse scale: further development and assessment of reliability and validity of a multidimensional partner abuse measure in clinical settings. *Violence and Victims, 20, 5, 529-547.*
- Hegarty, K., Sheehan, M., & Schonfeld, C. (1999). A multidimensional definition of partner abuse: development and preliminary validation of the Composite Abuse Scale. *Journal of Family Violence, 14, 4, 399-415.*
- Jones, N.T., Ji, P., Beck, M., & Beck, N. (2002). The reliability and validity of the revised Conflict Tactics Scale (CTS2) in a female incarcerated population. *Journal of Family Issues, 23, 3, 441-457.*
- Kimmel, M.S. (2002). "Gender symmetry" in domestic violence: A substantive and methodological research review. *Violence against women, 8, 11, 1332-1363.*
- Langhinrichsen-Rohling, J., Selwyn, C., & Rohling, M.L. (2012). Rates of bidirectional versus unidirectional intimate partner violence across samples, sexual orientations, and race/ethnicities: A comprehensive review. *Partner Abuse, 3, 2, 199-230.*
- Miller, E., Breslau, J., Petukhova, M., Fayyad, J., Green, J.G., Kola, L., Seedat, D.J., Stein, A., Viana, M.C., Andrade, L.H., Demyttenaere, K., de Girolamo, G., Haro, J.M., Hu, C., Kara, E.G., Kovess-Masfety, V., Tomov, T. & Kessler, R.C. (2011). Premarital mental disorders and physical violence in marriage: cross-national study of married couples. *The British Journal of Psychiatry, 199, 4, 330-337.*
- Mills, J.F., Kroner, D.G., & Forth, A.E. (1998). Novaco Anger Scale: reliability and validity within an adult criminal sample. *Assessment, 5, 237-248.*
- Murphy, C.M., & Hoover, S.A. (1999). Measuring emotional abuse in dating relationships as a multifactorial construct. *Violence and Victims, 14, 39-53.*

- Newman, B.S., & Campbell, C. (2011). Intimate partner violence among pregnant and parenting Latina adolescents. *Journal of Interpersonal Violence, 26*, 13, 2635-2657.
- O'Leary, K.D., & Williams, M.C. (2006). Agreement about acts of aggression in marriage. *Family Psychology, 20*, 656–662.
- Ramirez, I.L., Straus, M.A., 2006. The effect of question order on disclosure of intimate partner violence: An experimental test using the Conflict Tactics Scales. *Journal of Family Violence, 21*, 1–9.
- Reichenheim, M.E., Klein, R., & Moraes, C.L. (2007). Assessing the physical violence component of the Revised Conflict Tactics Scales when used in heterosexual couples: an item response theory analysis. *Cadernos de Saúde Pública, 23*, 1, 53-62.
- Relva, I.C., Fernandes, O.M., & Costa, R. (2013). Psychometric properties of revised conflict tactics scales: Portuguese sibling version (CTS2-SP). *Journal of Family Violence, 28*, 6, 577-585.
- Ro, E., & Lawrence, E. (2007). Comparing three measures of psychological aggression: Psychometric properties and differentiation from negative communication. *Journal of Family Violence, 22*, 7, 575-586.
- Ruehlman, L.S., & Karoly, P. (1991). With a little flak from my friends: Development and preliminary validation of the Test of Negative Social Exchange (TENSE). *Psychological Assessment, 3*, 97–104.
- Shepard, M.F. and Campbell, J.A. (1992). The Abusive Behaviour Inventory: A measure of psychological and physical abuse. *Journal of Interpersonal Violence, 7*, 3, 291-.305.
- Sherman, R., Fredman, N. (2013). *Handbook of measurements for marriage and family therapy*. London: Routledge.
- Sinha, M. (February, 2013). Measuring Violence Against Women: Statistical Trends. *Juristat: Component of Statistics Canada, 85-002-X*, 1-120.
- Straus, M.A. (1979). Measuring intrafamily conflict and violence: The conflict tactics (CT) scales. *Journal of Marriage and the Family, 75-88*.
- Straus, M.A. (1987). The Conflict Tactics Scales and its critics: an evaluation and new data on validity and reliability. *ERIC, ED297030*
- Straus, M.A. (1993). Identifying offenders in criminal justice research on domestic assault. *American Behavioral Scientist, 36*, 587-600.
- Straus, M.A. (1999). *Handbook for the conflict tactics scales (CTS)*. Durham: University of New Hampshire, Family Research Laboratory.
- Straus, M.A. (2004). Cross-cultural reliability and validity of the revised Conflict Tactics Scales: A study of university student dating couples in 17 nations. *Cross Cultural Research, 38*, 407-432.

- Straus, M.A. (2008). Dominance and symmetry in partner violence by male and female university students in 32 nations. *Children and Youth Services Review, 30, 3, 252-275.*
- Straus, M.A. (2010). Thirty years of denying the evidence on gender symmetry in partner violence: Implications for prevention and treatment. *Partner Abuse, 1, 3, 332-362.*
- Straus, M.A. (2015). Dyadic concordance and discordance in family violence: A powerful and practical approach to research and practice. *Aggression and Violent Behavior, 24, 83-94.*
- Straus, M.A., & Douglas, E.M. (2004). A short form of the Revised Conflict Tactics Scales, and typologies for severity and mutuality. *Violence and Victims, 19, 5, 507-520.*
- Straus, M.A., Gelles, R.J., & Steinmetz, S.K. (2006). *Behind closed doors: Violence in the American family* (2nd ed). NJ: Transaction Publications.
- Straus, M.A., Hamby, S.L., Boney-McCoy, S., & Sugarman, D.B. (1996). The revised conflict tactics scales (CTS2) development and preliminary psychometric data. *Journal of Family Issues, 17, 3, 283-316.*
- Straus, M.A., Hamby, S.L., & Warren, W.L. (2003). *The conflict tactics scales handbook: Revised conflict tactics scale (CTS2) and CTS: Parent-child version (CTSPC)*. Los Angeles, CA: Western Psychological Services.
- Tiwari, A., Fong, D.Y.T., Chan, K.L., Leung, W.C., Parker, B., & Ho, P.C. (2007). Identifying intimate partner violence: comparing the Chinese abuse assessment screen with the Chinese revised conflict tactics scales. *BJOG: An International Journal of Obstetrics & Gynaecology, 114, 9, 1065-1071.*
- Tjaden, P., & Thoennes, N. (2000). *Extent, nature, and consequence of intimate partner violence*. Washington, DC: U.S. Department of Justice.
- Winstok, Z. (2015). Critical review of Hamby's (2014) article titled "Intimate partner and sexual violence research, scientific progress, scientific challenges, and gender". *Trauma, Violence, & Abuse, DOI1524838015596962.*
- Winstok, Z., & Straus, M.A. (2016). Bridging the two sides of a 30-year controversy over gender differences in perpetration of physical partner violence. *Journal of Family Violence, 31, 8, 933-935.*
- Yan, E., & Tang, C. (2001). Prevalence and psychological impact of Chinese elder abuse. *Journal of Interpersonal Violence, 16, 11, 1158-1174.*
- Yun, S.H. (2011). Factor Structure and Reliability of the Revised Conflict Tactics Scales' (CTS2) 10-Factor Model in a Community-Based Female Sample. *Journal of Interpersonal Violence, 26, 4, 719-744.*