An Examination of Measurement Characteristics and Factorial Validity of the Revised Conflict Tactics Scale

Rae R. Newton, Cynthia Donaldson Connelly and John A. Landsverk
Educational and Psychological Measurement 2001 61: 317
DOI: 10.1177/0013164401612011

AN EXAMINATION OF MEASUREMENT CHARACTERISTICS AND FACTORIAL VALIDITY OF THE REVISED CONFLICT TACTICS SCALE

RAE R. NEWTON
California State University Fullerton

CYNTHIA DONALDSON CONNELLY
Children's Hospital and Health Center

JOHN A. LANDSVERK
San Diego State University

The present study investigated descriptive statistics for and factor validity of scores on the Revised Conflict Tactics Scale (CTS2) based on the responses of 295 high-risk postpartum women. The CTS2 yielded results for this sample similar to results obtained from a sample of college students in a previous study. Furthermore, confirmatory factor analytic results indicated that a five-factor model, specifying factors of Negotiation, Minor Psychological Aggression, Severe Psychological Aggression, Minor Physical Assault, and Severe Physical Assault, yielded better fit statistics than a three-factor model that combined the minor and severe scales within the domains of psychological aggression and physical assault.

The Conflict Tactics Scale (CTS) (Straus, 1979) is currently the most widely used measure of intimate partner violence (IPV); however, its use has not been without controversy (Dobash & Dobash, 1979; Pan, Neidig, & O'Leary, 1994; Ratner, 1998). In response to the criticism and recommendations, the Revised CTS (CTS2) was developed and, in principle, is a better measure of IPV than the widely used CTS. Although the preliminary psychometric findings are promising, further evidence concerning the measurement properties of this instrument is needed (Straus, Hamby, Boney-McCoy, & Sugarman, 1996). The present study is designed to examine the re-

Educational and Psychological Measurement, Vol. 61 No. 2, April 2001 317-335

liability and factorial validity of CTS2 scores among women. In particular, we (a) compared the psychometric findings based on a sample of college students, as reported by Straus et al. (1996), with our sample of high-risk postpartum women and (b) examined the factorial validity of CTS2 scores within this later sample.

The CTS (Straus, 1979, 1990) is designed to measure the extent to which partners in a dating, cohabiting, or marital relationship engage in psychological and physical attacks on each other and their degree of use of reasoning or negotiation to deal with conflicts (Straus et al., 1996). Conflict theory provides the theoretical basis for the CTS (Adams, 1965; Straus & Gelles, 1990; Straus, 1990). A major assumption of conflict theory is that all human association is effected by conflict in some way, and, without conflict, social groups fail to adapt to changed circumstance and lose their strength. Conflict, defined as the tactics or methods used to reach a resolution, is unavoidable and arises naturally from the quest for personal interests occurring at all levels of social functioning. The specific tactics employed to manage conflict may range from passive tactics, such as calm discussion, to forceful physical attacks. The CTS essentially operationalizes this range of conflict-related behaviors.

Numerous published studies have been based on data obtained by use of the CTS, including studies investigating the validity and reliability of CTS scores (reviewed in Straus, 1990). Previous researchers have used factor analytic procedures to establish the factor structure of CTS scores (Barling, O'Leary, Jouriles, Vivian, & MacEwen, 1987; Pan et al., 1994; Straus, 1990) and have suggested that IPV consists of three factors: Negotiation, Psychological Aggression, and Physical Assault. However, factor analysis of the CTS has also led to disagreements concerning which items should be included in the various scales and whether a distinction should be made between minor and severe levels of aggression (Barling et al., 1987).

In response to these and other criticisms and suggestions for revising and enlarging the CTS, the CTS2 was developed (Straus et al., 1996). The revision includes deletion of some items and inclusion of additional items in each of the three original scales, enabling more facets of each construct to be included and increasing reliability and validity of scale scores (Straus et al., 1996). For example, the original Reasoning scale has been replaced with an entirely new scale (Negotiation) composed of six new items that reflect actions taken to settle a conflict through discussion. Three of the items refer to cognitive aspects of negotiation and three to emotional aspects. Psychological Aggression, the second scale, includes items depicting the use of verbal and nonverbal acts that generally have the effect of being critical of or controlling the partner. Furthermore, the label *psychological aggression*

depicts a change from the original scale name (Verbal Aggression). This change was made because some of the acts, such as "stomped out of the room," are nonverbal aggressive acts (Straus et al., 1996). The Psychological Aggression scale retains four of the six original items, two in a slightly modified form, with the major change being the addition of four items. The third scale, Physical Assault, describes the acts that fit the definition of physical violence by a partner on which the CTS is based (Gelles & Straus, 1979; Straus et al., 1996). This third scale retains nine original acts with five slightly modified. The original item, "threatened him/her with a knife or gun," was dropped because this behavior was sufficiently covered by "used a knife or fired a gun," and three new items have been added (e.g., "slammed my partner against a wall"). Table 1 illustrates the changes as presented by Straus et al. (1996).

The CTS2 also provides a better operationalization of the distinction between minor and severe aggressive acts and applies this distinction to physical assault and psychological aggression. CTS2 examines an additional type of IPV (sexual coercion) and its consequence (physical injury from assaults by a partner). These two scales are considered collateral scales and are not further investigated in the present study. Thus, as described by Straus et al. (1996), the CTS2 may be represented by either a three-factor structure (Negotiation, Psychological Aggression, and Physical Assault) or a five-factor structure (Negotiation, Minor Psychological Aggression, Severe Psychological Aggression, Minor Physical Assault, and Severe Physical Assault).

The structure of the CTS2 has undergone empirical investigation among college students in which scores on the conflict tactics constructs demonstrated evidence of validity and reliability. Our literature review, however, found no published research that examines the factor structure of the revised CTS among other populations. Given this, structured methods are needed to understand the unobservable constructs that specify conflict tactics and to determine whether these latent factors remain consistent or differ across populations. The present study was undertaken to conduct psychometric analysis of the CTS2 within a cohort of high-risk postpartum women. We restricted our analyses to acts perpetrated on spouse or current partner by the female respondents. The specific research questions examined were: (a) How do the psychometric findings based on a sample of college students as reported by Straus et al. (1996) compare with our findings based on a sample of high-risk postpartum women? (b) Do three- and five-factor models of negotiation, psychological aggression, and physical assault depict adequate representations of the factor structure of the CTS2 with postpartum women? and (c) If so, does one of these alternative structures provide a better fit to the covariance matrix of the CTS2 items as represented by the present sample?

Table 1
Comparison of the Conflict Tactics Scale (CTS1) and the Revised Conflict Tactics Scale (CTS2)

CTS2		CTS1, Form R	
Revised Scale Name	Item	Original Scale Name	Item
Physical Assault	12	Violence	9
Minor	5		
Severe	7		
Psychological Aggression	8	Verbal Aggression	6
Minor	4		
Severe	4		
Negotiation	6	Reasoning	3
Total	26	Total	18
New scales		Additional item	
Injury	6	"Cried" (not scored)	1
Sexual Coercion	7		
Grand total	39	Grand total	19

Method

Sample and Procedure

The sample utilized in these analyses was a subset of 488 high-risk mothers, who gave birth from February 1996 to March 1997, participating in a randomized clinical trial of paraprofessional home visitation services (Healthy Families San Diego [HFSD]). Early Identification workers (EID) utilized two hospital-computerized systems that provided clinical and admission information to identify at-risk families. High-risk mothers were identified by a number of characteristics, including single-parent status, teen pregnancy, late or no prenatal care, low maternal educational attainment, lack of employment, family violence, parental substance abuse, and/or mental health problems.

Extensive demographic, lifestyle, and adaptation data were obtained using standardized measures. Annual interviews and assessments of the mothers, children, and the home environment were conducted through the first 3 years of each child's life. Additionally, all families received short telephone tracking surveys at 4, 8, 16, 20, 28, and 32 months. The sample was characterized by marked racial/ethnic diversity, and measures were used to assure reasonable confidence of comparability across language groups. As almost 20% of the HFSD sample stated a preference for using Spanish, several Spanish-speaking interviewers were employed to conduct interviews. The questionnaire and observations were conducted annually in families'

homes by trained research interviewers who were blinded to the families' group status in the study and were not connected with the home visitation component. At the end of the 1st year, the study retention rate was 90%.

Of the 488 postpartum women, 90% (435) completed interviews at Year 1. Participants ranged in age from 14 to 42 years with a mean age of 23.39 (*SD* = 6.19). Of these, 45.6% were 21 years or younger. Over half of the mothers (54.3%) had not completed high school, and another 21.8% had no education beyond high school graduation or the general equivalency diploma (GED). Although single motherhood was dominant with 342 (78.6%) not married, 50.8% of the mothers reported the father of the baby (FOB) as current partner and 54.9% reported seeing the baby's father daily. The sample was ethnically diverse: 118 (27.1%) were Hispanic/English, 83 (19.1%) were Hispanic/Spanish, 84 (19.3%) were African American, 109 (25.1%) were Caucasian, and 41 (9.4%) were Asian, Pacific Islander, Native American, or other.

For the purpose of our analyses, data collected from the participants using the CTS2 were of interest; therefore, a subset of the 435 women was sampled based on the following inclusion criteria: (a) completion of the CTS2, (b) having a current partner, and (c) missing no more than two CTS2 items. Of the 295 meeting these criteria, 256 had complete data, 33 were missing one item, and 6 were missing two items. Of the 6 missing two items, only 1 was missing items from the same scale (Negotiation). Missing values were replaced using the mean of the scale for which the item was missing, based on the five-scale structure.

Instrumentation

CTS2. The CTS2 is a 39-item self-report instrument designed to measure the extent to which partners in a dating, cohabiting, or marital relationship have engaged in psychological and physical attacks on each other and the extent to which the partners have used reasoning or negotiation to deal with conflicts during the past 12 months (Straus et al., 1996). The 39 items comprise the five major scales of the CTS2: Negotiation, Psychological Aggression, Physical Assault, Sexual Coercion, and Physical Injury From Partner Assaults. The present study focused on the 22 items that compose the Negotiation, Psychological Aggression, and Physical Assault scales. Thus, reference to either a three-factor model or a five-factor model is based on a five-factor model that divides the Psychological Aggression and Physical Assault scales into minor and severe subscales (not the difference between models with and without the Sexual Coercion and Physical Injury scales). Items composing these collateral scales (coercion and physical injury) are unavoidably confounded with the Psychological Aggression and Physical Assault scales and could not logically be included in our test of a congeneric measurement model.

The CTS2 replaces the matrix format of the CTS1 (interview schedule) with a format developed by Neidig (1990). Each item addresses what the participant has done and then is repeated on the next line for what the partner has done. The hierarchical order used in the CTS1 has also been replaced with an interspersed order in the CTS2. Because the present study focused on the structure of female patterns of conflict, we examined only the first question in each pair, assessing what the female participant reported that she had done to her partner.

The Negotiation scale is a six-item scale defined as actions taken to settle a disagreement through discussion. This scale can be divided into Cognitive and Emotional subscales. The items in the Cognitive subscale are examples of discussion foci (e.g., suggested a compromise to a disagreement). The Emotion subscale is meant to measure the extent to which positive affect is communicated by asking about expression of feelings of care and respect for the partner (e.g., "I showed my partner I cared even though we disagreed"). Psychological Aggression is the new name for the Verbal Aggression scale. This change was made because some of the acts are nonverbal aggressive acts, such as "stomped out of room" (Straus et al., 1996). The Physical Assault scale replaced the Violence scale because this verbiage better describes the acts that fit the definition of physical violence and thus avoids confusion with the use of violence. Straus et al. (1996) reported internal consistency coefficients for scores on the CTS2 scales as high as or higher than coefficients for scores on the CTS1 scales.

Scoring the CTS2 follows the principles used to score the CTS1 (see Straus et al., 1996). Response categories range from 0 (*never*) to 6 (*more than 20 times*) for each of the tactics with the referent period being the past 12 months. Straus et al. (1996) suggested a strategy for recoding response categories 3 to 6 to reflect the midpoints of the category label. For category 3 (*3 to 5 times*) the midpoint is 4, for category 4 (*6 to 10 times*) it is 8, for category 5 (*11 to 20 times*) it is 15, and for category 6 (*more than 20 times in the past year*) Straus recommended using 25 as the midpoint (see Straus et al., 1996, p. 305). For most of the CTS2 items, particularly those with infrequent high scores, this recoding scheme adds considerably to the skew and kurtosis of the data and, accordingly, creates problems for statistical models built on parametric procedures and/or assumptions of multivariate normality. Thus, there may be some rationale for not instituting these recodes prior to certain types of analyses (Straus, personal communication, 1999).

Results

Table 2 presents descriptive statistics and Cronbach's alpha reliability coefficients for our sample of 295 women. Also included are alpha coefficients for 317 college students as reported by Straus et al. (1996). The Straus et al. reliability coefficients include male and female respondents and are

Table 2
Descriptive Statistics and Reliability Coefficients for Scores on the Revised Conflict Tactics Scales (CTS2)

Scale	Number of Items	Minimum	Maximum		*M*	*SD*	Alpha$_1$[a]	Alpha$_2$[b]	Alpha$_3$[c]
			Potential	Observed					
Negotiation	6	0	150	150	56.5	41.3	.84	.86	.86
Psychological Aggression									
Total	8	0	200	124	19.7	23.8	.70	.77	.79
Minor	4	0	100	90	16.6	19.1	.66	.74	NA
Severe	4	0	100	46	3.1	7.7	.49	.63	NA
Physical Assault									
Total	12	0	300	69	4.2	10.1	.67	.78	.86
Minor	5	0	125	52	3.3	8.3	.62	.75	NA
Severe	7	0	175	32	.9	3.0	.43	.57	NA

Note. $N = 295$ for all statistics except Alpha$_3$.
a. Alpha computed after applying Straus, Hamby, Boney-McCoy, and Sugarman (1996) recode structure.
b. No recode prior to computing alpha.
c. Alphas reported by Straus et al. (1996). $N = 317$.

reported as approximations to the final version of the CTS2. Coefficients for minor and severe psychological aggression and minor and severe assault were not reported by Straus et al. and are thus not included in Table 2. Note that all statistics, except $Alpha_2$, are based on the CTS2 items after applying the scoring recode suggested by Straus et al.

Table 2 shows that only the Negotiation scale spanned the potential range of scores obtainable on these scales. Our sample reports an average of 56.6 attempts to negotiate with partner by the female respondents during the past year. Scales for Psychological Aggression and Physical Assault, as would be expected, have both lower mean scores and less variability. This sample experienced an average of 19.7 acts of Psychological Aggression (16.6 Minor and 3.1 Severe) and 4.2 acts of Physical Assault (3.3 Minor and .9 Severe). Note that when comparing means for Minor versus Severe Psychological Aggression and Physical Assault, the Severe subscale produces a much smaller mean score even though, as in the case of Severe Physical Assault, the potential score is greater.

Table 2 also contains three columns of alpha reliability coefficients based on the original scores and the recoded scores, as well as scores for the Minor and Severe subscales of Psychological Aggression and Physical Assault. Recoding, though presenting a more realistic picture of the actual frequency of occurrence of the behaviors in question, also seriously skews these distributions and in our sample produces attenuated coefficients for some scales.

The alpha coefficients for scores on Negotiation are highly similar and suggest that this scale (like its predecessor, Reasoning) produces internally consistent scores. Similarly, the alpha coefficient for scores on Psychological Aggression, using the original scores, is almost identical to that reported by Strauss et al. (1996) for the Psychological Aggression approximation, and both are close to the .80 rule of thumb considered desirable (.77 and .79). Alpha using the recoded scores is much lower (.70), suggesting that the skew created by executing this recode attenuates the reliability of the scores. We observed similar attenuation in all subsequent scale scores. The greatest divergence between our observed reliability coefficients and those reported by Straus et al. is evidenced in the Physical Assault scale scores. Whereas Straus et al. reported a value of .86, we observed .78 for the original scores and .67 for the recoded scores. Clearly, the data for Straus et al.'s coed group of college students produced higher reliabilities than our sample of poorly educated women.

When breaking the Psychological Aggression and Physical Assault scales into Minor and Severe subscales, we observed lower reliabilities for the Severe subscales in both instances, even though the Severe subscale of the Physical Assault scale contains more items than its minor counterpart (7 vs. 5). We would generally expect scales containing fewer items to produce scores with lower reliability coefficients, other things being equal; however,

it appears that a greater proportion of the unreliability in the scores on the Psychological Aggression and Physical Assault scales is attributable to the items composing the severe behaviors of these scales.

Prevalence and Chronicity

Prevalence refers to the percentage of the sample that reports at least one instance of the behavior in question. Chronicity refers to the frequency of occurrence of behaviors among those who report at least one instance of the behavior in question. Thus, for each scale and subscale, the chronicity mean is based on only those who were coded as reporting at least one instance of the behavior, and thus the sample size varies for each scale. These statistics are reported in Table 3 for both the present study and the Strauss et al. (1996) study, where available. We expected the prevalence and chronicity rates in our sample to be higher than those reported by Straus et al., except for Negotiation, where we expected the rates to be lower. Contrary to our expectation, the prevalence rates for Negotiation and Psychological Aggression were nearly identical in both samples. As expected, prevalence rates for Physical Assault were higher. Also, as expected, the chronicity mean for Negotiation was higher for college women (almost 70%) than for the present sample (about 59%). Similarly, the chronicity mean for Psychological Aggression was almost 8% higher for the present sample than for the sample of college women (23.6% vs. 16%, respectively). An unexpected finding was the similarly high chronicity means of Physical Assault for both groups of women (9.6 for our sample and 9.4 for college students), suggesting that among those women for whom physical assault is seen as an option, it occurs frequently. It was not possible to examine the rates of Minor versus Severe Assault in the Straus et al. report; however, for the 18.6% of the present sample who reported at least one instance of severe physical assault on their partner, the average number of severe assaults was 4.6 in the past year.

Table 4 presents the Pearson product moment and Spearman correlation coefficients for our data and the Pearson correlations for the Straus et al. (1996) data for Negotiation, Psychological Aggression, and Physical Assault scales. The expected low correlations between Negotiation and Physical Assault are highly similar to those reported by Straus et al.

Examining the Factorial Validity of the CTS2

Research Questions and Guiding Assumptions

Our second major goal was to examine the factorial validity of the CTS2 scores. As presented by Straus et al. (1996), the CTS2 is composed of three major factors, with individual items defined as representing minor and severe

Table 3
Prevalence and Chronicity Statistics

Scale	Number of Items	Newton, Connelly, and Landsverk[a] Prevalence (%)	Newton, Connelly, and Landsverk[a] Chronicity (*M*)	Straus, Hamby, Boney-McCoy, and Sugarman (1996)[b] Prevalence (%)	Straus, Hamby, Boney-McCoy, and Sugarman (1996)[b] Chronicity (*M*)
Negotiation	6	95.6	59.1	98	69.7
Psychological Aggression					
Total	8	83.4	23.6	83	16.0
Minor	4	82.7	20.1	NA	NA
Severe	4	35.9	8.5	NA	NA
Physical Assault					
Total	12	43.1	9.6	35	9.4
Minor	5	40.7	8.1	NA	NA
Severe	7	18.6	4.6	NA	NA

Note. Data are for women only.
a. $N = 295$ for all prevalence statistics; N varies for chronicity means. To find the chronicity N, multiply the prevalence proportion by 295.
b. $N = 204$ for all prevalence statistics; N varies for chronicity means. To find the chronicity N, multiply the prevalence proportion by 204.

Table 4
Parametric and Nonparametric Correlations of the Revised Conflict Tactics Scales (CTS2)

Scale	Negotiation	Psychological Aggression
Psychological Aggression		
Pearson	.372*	1.00
Spearman	.458*	1.00
Straus et al.[a]	.40*	1.00
Physical Assault		
Pearson	.159*	.541*
Spearman	.203*	.526*
Straus et al.[a]	.21*	.67*

Note. $N = 295$.
a. Reported in Straus, Hamby, Boney-McCoy, and Sugarman (1996); $N = 204$.
*$p < .001$.

psychological aggression and physical assault. Our main, two-part question was, Do the three- and five-factor models depict adequate representations of the factor structure of the CTS2, and, if so, does one of these alternative structures provide a better fit to the covariance matrix of the CTS2 items? Our interest in a five-factor solution rests on the assumption that such a model, should it be empirically supported, would permit more definitive assessment

of the nature of conflict within different populations. This is not a unique consideration. Research with the Violence scale of the CTS1 has addressed a similar question (Hornung, McCullough, & Sugimoto, 1981; Pan et al., 1994; Straus, 1979). Pan et al. (1994) concluded that it was reasonable to divide the CTS1 Violence scale into minor and severe factors but did not base this conclusion on confirmatory factor analysis (CFA).

To answer these questions, CFA utilizing maximum likelihood estimation methods with the EQS computer program (Bentler & Wu, 1995) was employed. Following the work of Straus et al. (1996) and our desire to focus on the factorial validity of the CTS2 scores, we made a number of a priori assumptions and decisions that guide these analyses: (a) responses to the CTS2 could be explained by either a three- or five-factor first-order model; (b) for the three- and five-factor solutions, each item would be associated with only the factor it was designed to measure, and all other coefficients would be fixed to zero; (c) all factors would be allowed to covary to allow for an oblique factor model; (d) post hoc model fitting would be kept to a minimum and would exclude the addition of factorially complex items; and (e) correlated error terms would be restricted within latent constructs (i.e., factors). Improvement in fit between models with and without correlated error terms would be assessed using the chi-square difference test. This test is appropriate for the assessment of the statistical significance of difference in fit between nested models. Note that the three- and five-factor versions of the CTS2 represent nonnested models, but the significance of the difference between versions of either of these models (i.e., with and without the inclusion of correlated errors) does represented nested models. Thus, the difference in fit between models with and without correlated error terms can be tested with the chi-square goodness-of-fit test, but the difference between three- and five-factor models cannot be tested in this manner.

A decision needed to be made regarding whether to analyze the covariance matrix represented by the original data codes (with scores ranging from 0 to 6) or the covariance matrix based on the recoded scores (ranging from 0 to 25). Although both matrices include highly kurtotic, positively skewed distributions of some CTS2 items, the imposition of a scoring system that greatly expands the magnitude of these indicators of nonnormality seemed unwarranted. For example, the skew of the "used a knife or gun on my partner" item for the recoded data was 16 ($SE = .142$), and the kurtosis was 277 ($SE = .283$). The typical fix for such a positively skewed distribution would be to institute a square root or other transformation of the original scores (Box & Cox, 1964; Mosteller & Tukey, 1977). However, it made little sense to first exacerbate the positive skew of some items by imposing the scoring system described by Straus and colleagues (Straus & Gelles, 1990; Straus, 1995) and then undoing this by transforming the distribution. Thus, we found use of the

data in its original form (i.e., scores ranging from 0 to 6) to be the most reasonable approach, and all analyses reported herein are based on the covariance matrix derived from these original (i.e., unrecoded) scores.

Finally, given the known skew and kurtosis of some CTS2 items, the Satorra-Bentler scaled chi-square statistic (S-Bχ^2) and its associated fit index, the robust comparative fit index (RCFI), were used (Satorra & Bentler, 1988). S-Bχ^2 uses a scaling correction for the chi-square statistic when distributional assumptions are unwarranted. Byrne, Baron, and Campbell (1994) argued that the RCFI should be the measure of choice when the Satorra-Bentler scaling correction for the χ^2 (S-Bχ^2) statistic is warranted, as "it allows for a more cogent assessment of factorial validity than is possible with the uncorrected (i.e., biased) statistic" (p. 171). In addition, assessment of fit included comparative and parsimonious fit indices. Comparative fit indices involve comparison with a baseline model. (The most common baseline model is the null or independence model that specifies no relationships between the variables composing the model.) Parsimonious fit indices are concerned primarily with the cost-benefit tradeoff of fit and degrees of freedom. These measures are relevant in this case because we wished to compare two nonnested models, one of which is more factorially complex than the other. Baseline three- and five-factor models are presented in Figure 1.

Confirmatory Factor Analysis of Three- and Five-Factor Models

Table 5 presents the fit statistics for three- and five-factor models with and without correlated error terms. A number of conclusions can be reached by examination of this table. First, none of the models that exclude correlated error terms fit particularly well. This might be expected in light of the restrictive nature of these models and the small number of parameters being used to establish fit. Given these findings, we respecified the model to include additional parameters identified using the Lagrange Multiplier Test and our a priori criteria for adding parameters to the model. It should be reemphasized that our goal was not to simply improve the fit of the model by relaxing model constraints but to assess the factorial validity of the CTS2 scores. Thus, our respecification of the model disallowed factorially complex items and error correlations that cross the five factors. In addition, the decision was made that any parameter estimated in one model would also be estimated in the remaining model. Using these criteria, four error covariances were estimated: one on the Negotiation factor, two on the Minor Physical Assault factor, and one on the Severe Physical Assault factor. For the Negotiation factor, the items were "suggested a compromise" with "agreed to try a solution." For the Minor Physical Assault factor, the items were (a) "threw something at my partner" with "I twisted my partner's arm or hair" and (b) "I grabbed my partner" with

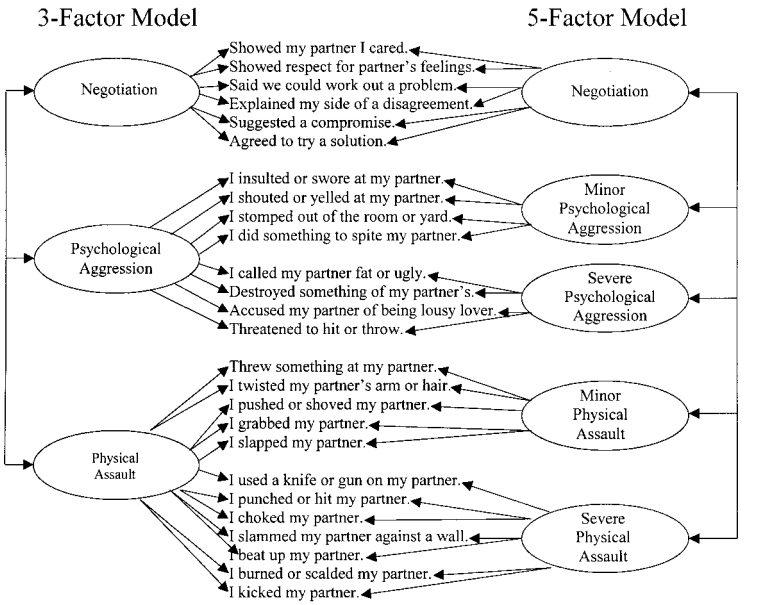

Figure 1. Baseline three- and five-factor Revised Conflict Tactics Scale (CTS2) confirmatory factor analysis models.

"I slapped my partner." For the Severe Physical Assault factor, the items were "I choked my partner" with "I burned or scalded my partner." The face validity of the connection between these items is evident and, as shown in Table 5, the improvement in fit is noteworthy, with the chi-square difference ranging from a high of 109 ($df = 4$, $p < .001$) for the three-factor model to a low of 94 ($df = 4$, $p < .001$) for the five-factor model. The decision as to when model fitting is complete is a difficult one and must be based on statistical and conceptual considerations. Despite our rather prohibitive guidelines for including additional model parameters, there were a number of additional parameters that could have been estimated. Our goal, however, was to offer a restrictive test of the factorial validity of the CTS2, not to demonstrate the best fitting of all possible models.

Regardless of which measure of fit one chooses, the five-factor solution consistently fit better than the three-factor solution. Fit statistics for the five-factor solutions ranged from a Bentler-Bonnet Normed Fit Index (NFI) of .785 to an RCFI of .922. Despite the evidence that the RCFI based on the five-factor solution surpasses the .90 value generally considered the cutoff point for well-fitting models, the variability in the various fit indices suggests evidence of misfit in this model. Some of this lack of fit is clearly attributable to restrictions placed on error covariances; however, examina-

Table 5
*Revised Conflict Tactics Scale (CTS2) Fit Statistics: Three- and Five-Factor Oblique Models, Maximum Likelihood Solutions (*df *in parentheses)*

	Three-Factor Solution		Five-Factor Solution	
	No Correlated Errors	Four Correlated Errors	No Correlated Errors	Four Correlated Errors
Independence model chi-square	2,988 (325)	3,062 (325)	2,988 (325)	3,062 (325)
Model chi-square	1,459.7 (296)	804.1 (292)	1,316.0 (289)	657.7 (285)
Satorra-Bentler scaled chi-square	544.3 (296)	434.7 (292)	457.2 (289)	363.0 (285)
AIC independence model	2,338	2,412	2,338	2,412
AIC	868	220	738	88
CAIC independence model	814	889	814	889
CAIC	–520	–1,149	–617	–1,248
Bentler-Bonett Normed Fit Index	.511	.737	.560	.785
Bentler-Bonett Nonnormed Fit Index	.520	.792	.566	.845
Comparative Fit Index	.562	.813	.614	.864
Robust Comparative Fit Index	.495	.857	.658	.922

Note. AIC = Akaike's Information Criterion, CAIC = Consistent Akaike's Information Criterion. $N = 295$ for all analyses. Chi-square difference tests for models with and without correlated errors: three-factor, $\chi^2 = 109.3$, $df = 4$, $p < .001$; five-factor, $\chi^2 = 94.2$, $df = 4$, $p < .001$.

tion of the factor parameter estimates and corresponding R^2 values for the three- and five-factor solutions also suggests additional areas of misfit. These values are shown in Table 6. An examination of the factor parameters for the three- and five-factor models shows that the first 19 items all had coefficients of .4 or above, with the exception of Item 13, which is clearly different in item content from the other items. Item 13 ("accused my partner of being a lousy lover") produced coefficients of .352 and .361 for the three- and five-factor solutions, respectively.

The three- and five-factor solutions included estimation of the correlations between factors. These relationships provide valuable information concerning CTS2. First, it is important to note that both solutions replicate earlier findings reported in Table 4. The correlation between Psychological Aggression and Physical Assault in the three-factor solution is .813, and the correlation between Negotiation and Physical Assault is .258. In the five-factor model, we were able to break out the distinctions between the Minor and Severe subscales. As the severity of Psychological Aggression and Physical Assault increased, the relationship of these scales' scores with Negotiation decreased. Second, Minor Psychological Aggression was more highly correlated with Minor Physical Assault than Severe Physical Assault, and Severe Psychological Aggression was more highly correlated with Severe Physical Assault than Minor Physical Assault.

Discussion

The above analyses permit tentative answers to a number of questions concerning the generalizability, psychometric properties, and practical utility of the CTS2. In addition, a number of questions about specific subscales are raised that will require additional research. Our results, taken in toto, suggest that the CTS2 functioned well, but not as well as reported by Straus with a coed college student population. Straus reported Cronbach alpha coefficients of .86, .79, and .86 for Negotiation, Psychological Aggression, and Physical Assault, respectively. Although we replicated the findings for the Negotiation scale, our reliability coefficients were slightly lower for scores on the Psychological Aggression and Physical Assault scales. The degree to which our coefficients can be considered lower, and whether this can be considered a substantively important difference, depends on how the data were treated prior to calculating the alpha coefficient. Our reliability coefficients were much closer to the .8 standard when using the original data codes, .77 for Psychological Aggression and .78 for Physical Assault. These were reduced to .70 and .67, respectively, when using the recoded data. There is no a priori explanation for this attenuation in value, as extreme data points might be expected to exaggerate a correlation as well as shrink it. Further exploratory data analysis needs to be undertaken to discover the causes for this attenuation in the coefficient values.

Table 6
Revised Conflict Tactics Scale (CTS2) Confirmatory Factor Analysis: Maximum Likelihood Estimates, Explained Variance, and Factor and Error Correlations for Three- and Five-Factor Models

	Three-Factor Solution		Five-Factor Solution	
	Loading	R^2	Loading	R^2
Negotiation (three- and five-factor)				
1. Showed my partner I cared.	.738	.545	.738	.545
2. Showed respect for partner's feelings.	.661	.437	.659	.435
3. Said we could work out a problem.	.671	.451	.668	.446
4. Explained my side of a disagreement.	.819	.671	.821	.674
→ 5. Suggested a compromise.	.646	.418	.648	.420
→ 6. Agreed to try a solution.	.640	.410	.641	.411
Psychological Aggression (three-factor)				
Minor (five-factor)				
7. I insulted or swore at my partner.	.628	.394	.737	.544
8. I shouted or yelled at my partner.	.693	.481	.795	.623
9. I stomped out of the room or yard.	.563	.317	.633	.401
10. I did something to spite my partner.	.460	.212	.423	.179
Severe (five-factor)				
11. I called my partner fat or ugly.	.454	.206	.448	.201
12. Destroyed something of my partner's.	.568	.322	.608	.370
13. Accused my partner of being lousy lover.	.352	.124	.361	.131
14. Threatened to hit or throw.	.694	.482	.755	.571
Physical Assault (three-factor)				
Minor (five-factor)				
→15. Threw something at my partner.	.687	.472	.686	.470
→16. I twisted my partner's arm or hair.	.411	.169	.399	.159
17. I pushed or shoved my partner.	.814	.663	.821	.674
→18. I grabbed my partner.	.591	.350	.608	.370
→19. I slapped my partner.	.425	.181	.432	.187
Severe (five-factor)				
20. I used a knife or gun on my partner.	.184[a]	.034	.174	.030
21. I punched or hit my partner.	.660	.436	.679	.460
→22. I choked my partner	.166[a]	.027	.178[a]	.032
23. I slammed my partner against a wall.	.248	.062	.260[a]	.068
24. I beat up my partner.	.173	.030	.158[a]	.025
→25. I burned or scalded my partner.	.124[a]	.015	.149[a]	.022
26. I kicked my partner.	.665	.442	.729[a]	.531

	Three-Factor Solution			Five-Factor Solution				
	1	2	3	A	B	C	D	E
Factor correlations								
Negotiation (1, A)	1.0			1.0				
Psychological Aggression (2)	.525	1.0						
Minor (B)				.598	1.0			
Severe (C)				.319	.692	1.0		

Table 6 Continued

	Three-Factor Solution			Five-Factor Solution				
	1	2	3	A	B	C	D	E
Factor correlations								
Physical Assault (3)	.258	.813	1.0					
Minor (D)				.285	.638	.905	1.0	
Severe (E)				.192[a]	.445[a]	.982	.906	1.0
Error term correlations								
→ 5. Suggested a compromise. → 6. Agreed to try a solution.	.462			.461				
→15. Threw something at my partner. →16. I twisted my partner's arm or hair.	.334[a]			.345[a]				
→18. I grabbed my partner. →19. I slapped my partner.	.321[a]			.311				
→22. I choked my partner. →25. I burned or scalded my partner.	.665[a]			.662[a]				

Note. $N = 295$ for all analyses. → indicates correlated error terms.
a. Not statistically significant using robust standard errors.

Our analysis also included the alpha values for scores on the Minor and Severe subscales of Psychological Aggression and Physical Assault. As would be expected, the attenuation in the values of these coefficients was also observed when the Straus recoding system was applied. One would also expect that, because the number of items in each subscale is smaller, the value of Cronbach's alpha would be smaller; however, for the Minor Psychological Aggression and Minor Physical Assault subscales, these reductions are identical and quite small (.03), even though the number of items in the respective subscale is half that of the full scale. Although one would not expect scores from four- or five-item scales to produce high reliability coefficients, the differences between the Minor and Severe subscales of Psychological Aggression and Physical Assault suggest that the Minor subscales fit together as a unit better than the Severe subscales.

Prevalence and chronicity statistics in the two samples produced results that were predictable, given the nature of the two populations from which the samples came, but also produced some disconcerting findings. First, based on our a priori knowledge of the two samples (one of college students, the other of poorly educated high-risk mothers), we expected that the prevalence and chronicity of Psychological Aggression and Physical Assault would be higher among the high-risk sample. In fact, the prevalence of Psychological

Aggression was virtually identical in the two samples (83%), and the chronicity of Psychological Aggression was one third higher in the high-risk sample. Examination of the subscales in our sample indicated higher chronicity of Minor than Severe Psychological Aggression and greater evidence of Physical Assault than in the Straus college sample. Nevertheless, the chronicity statistics for Physical Assault were virtually identical. The one in three college women in Straus's sample who assaulted their partner did so an average of 9.4 times, whereas the two in five women in the present sample who assaulted their partner did so an average of 9.6 times, a negligible difference. It would be interesting to compare the prevalence and chronicity of Minor and Severe forms of Psychological Aggression and Physical Assault in the two samples of women, but these data were not reported by Straus et al. (1996).

Finally, we considered the question of the factorial validity of the CTS2 scores. The feasibility of three- and five-factor solutions was examined. Models without correlated errors did not fit the data well. However, the relaxation of only four such constraints improved the fit of three- and five-factor models appreciably. Also, five-factor models always fit better than three-factor models. Moreover, only one measure of fit, the RCFI, with only one model, the five-factor model with correlated errors, met the generally accepted criterion of a good fit of .9 or higher, suggesting the need for additional evaluation of individual model parameters.

The source of lack of fit was locatable within the seven items composing the Severe subscale of Physical Assault. With only a few exceptions, the remaining factor parameters and associated R^2 values indicated substantial amounts of explained variance. However, only two of the seven items of the Severe Assault subscale explain more than 7% of the variance. This is surprising, given that Straus et al. (1996) report item-total correlations ranging from .39 to .70 for these items. Perhaps the difference lies in the fact that not all severe acts of violence occur together but do occur in concert with other types of physical assault. If this were true, we would expect the three-factor model to fit better than the five-factor model; however, the statistics for the model are highly similar.

References

Adams, B. N. (1965). Coercion & consensus theories: Some unresolved issues. *American Journal of Sociology, 71*, 714-716.

Barling, J., O'Leary, K., Jouriles, E., Vivian, N., & MacEwen, K. (1987). Factor similarity of the Conflict Tactics Scales across samples, spouses, and sites: Issues and implication. *Journal of Family Violence*, *2*, 37-55.

Bentler, P. M., & Wu, E.J.C. (1995). *EQS/Windows user's guide: Version 5.7b.* Los Angeles: BMDP Statistical Software.

Box, G.E.P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, 26*, 211-243.

Byrne, B., Baron, P., & Campbell, T. (1994). The Beck Depression Inventory (French version): Testing for gender-invariant factorial structure for nonclinical adolescents. *Journal of Adolescent Research*, *9*(2), 166-179.

Dobash, R. E., & Dobash, R. P. (1979). *Violence against wives: A case against patriarchy*. New York: Free Press.

Gelles, R. J., & Straus, M. A. (1979). Determinants of violence in the family: Toward a theoretical integration. In W. R. Burr, R. Hill, F. I. Nye, & I. L. Reiss (Eds.), *Contemporary theories about the family* (Vol. 1, pp. 549-581). New York: Free Press.

Hornung, C., McCullough, B., & Sugimoto, T. (1981). Status relationship in marriage: Risk factors in spouse abuse. *Journal of Marriage and the Family*, *43*, 675-692.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression*. Reading, MA: Addison-Wesley.

Neidig, P. (1990). *A modification of the Conflict Tactics Scale*. Stonybrook: Department of Psychology, State University of New York at Stonybrook.

Pan, H. S., Neidig, P. H., & O'Leary, K. D. (1994). Male-female & aggression-victim differences in the factor structure of the modified Conflict Tactics Scales. *Journal of Interpersonal Violence*, *9*, 366-382.

Ratner, P. (1998). Modeling acts of aggression and dominance as wife abuse and exploring their adverse health effects. *Journal of Marriage and the Family*, *60*, 452-465.

Satorra, A., & Bentler, P. (1988). *Scaling corrections for chi-square statistics in covariance structure analysis: Proceedings of the Business and Economics Sections*. Alexandria, VA: American Statistical Association.

Straus, M., & Gelles, R. (Eds.). (1990). *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families*. New Brunswick, NJ: Transaction Press.

Straus, M., Hamby, S., Boney-McCoy, S., & Sugarman, D. (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*, *17*(3), 283-316.

Straus, M. A. (1979). Measuring intrafamily conflict and violence. The Conflict Tactic (CT) Scales. *Journal of Marriage and the Family*, *41*, 75-88.

Straus, M. A. (1990). The Conflict Tactics Scales and its critics: An evaluation and new data on validity and reliability. In M. Straus & R. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 49-73). New Brunswick, NJ: Transaction Press.

Straus, M. A. (1995). *Manual for the Conflict Tactics Scales*. Durham: Family Research Laboratory, University of New Hampshire.