Cross-Cultural Research
http://ccr.sagepub.com/

Cross-Cultural Reliability and Validity of the Revised Conflict Tactics Scales: A Study of University Student Dating Couples in 17 Nations

Murray A. Straus

Cross-Cultural Research 2004 38: 407
DOI: 10.1177/1069397104269543

The online version of this article can be found at:
http://ccr.sagepub.com/content/38/4/407

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Society for Cross-Cultural Research
Sponsored by the Human Relations Area Files

Additional services and information for *Cross-Cultural Research* can be found at:

Email Alerts: http://ccr.sagepub.com/cgi/alerts

Subscriptions: http://ccr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://ccr.sagepub.com/content/38/4/407.refs.html

>> Version of Record - Sep 25, 2004

What is This?

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

Cross-Cultural Reliability and Validity of the Revised Conflict Tactics Scales: A Study of University Student Dating Couples in 17 Nations

Murray A. Straus
University of New Hampshire

Although the original Conflict Tactics Scales (CTS) have been successfully used in many countries, no studies have been published on the cross-cultural reliability and validity of the revised instrument (CTS2). This study is intended to provide some of the needed psychometric information. It reports reliability and examples of validity evidence for the five CTS2 (physical assault, physical injury, psychological aggression, sexual coercion, and negotiation) to measure these aspects of the dating relationships of 7,179 students at 33 universities in 17 countries. The results show high alpha coefficients of internal consistency and low confounding with social desirability response set. Examples indicating the construct validity of the CTS2 Physical Assault and Injury Scales are also presented. Although the data refer to dating relationships of university students, the results are sufficiently promising to encourage use of the CTS2 in a variety of cultural settings.

Cross-Cultural Research, Vol. 38 No. 4, November 2004 407-432
DOI: 10.1177/1069397104269543
© 2004 Sage Publications

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

Keywords: *aggression; crime; cross-cultural; injury; measure; negotiation; psychological; reliability; sex; validity; violence*

Many books and hundreds of articles on violence between partners who are dating, cohabiting, and married have used data provided by the original Conflict Tactics Scales (CTS; Straus, 1990b). The revised CTS (CTS2; Straus, Hamby, Boney-McCoy, & Sugarman, 1996) include improved versions of the three original scales that measure physical assault, psychological aggression, and negotiation and add two supplemental scales to measure injury resulting from an assault by a partner and sexual coercion.

There is a large amount of research showing that the CTS have a stable factor structure and moderate to high reliability (Archer, 1999; Yodanis, Hill, & Straus, 1997). There is also extensive evidence of construct validity (Straus, 1990a). The original CTS have been successfully used in many countries and with different ethnic groups within the United States (Yodanis et al., 1997). However, the CTS were revised in 1996 (Straus et al., 1996) to improve question wording, to add items to provide a more detail coverage of severe assaults, and to add scales to measure injury and sexual coercion. These changes raise the question of whether the evidence of reliability, validity, and cross-cultural applicability of the original CTS also apply to the CTS2.

The CTS2 have been used much less often because they are relatively recent and probably also because potential users learn about the CTS from the many articles and books using the original CTS that of course do not refer to the revised version. In addition, the

---

Author's Note: *Other articles on this and related issues can be downloaded from the Web site http://pubpages/~mas2. This article is part of the International Dating Violence Study (IDVS). The IDVS is being conducted by a consortium of researchers who obtained the data for their country. I am grateful to the IDVS consortium members for allowing me to use their data for the purposes of this article. A list of the consortium members and their addresses are available on the previously mentioned Web site. It is also a pleasure to express my appreciation to Sarah Savage and Christy Knox for the statistical analysis and preparation of the tables. Financial support has been provided by the University of New Hampshire and the National Institute of Mental Health grant T32MH15161. For purposes of compliance with Section 507 of P.L. 104-208, readers are advised that 25% of the funds for this project are derived from federal sources. The total amount of federal funding involved is $56,632.*

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

CTS2 are much longer because they include supplemental scales to measure physical injury and sexual coercion, and potential users may not realize that these two supplemental scales can be dropped to provide an instrument of equivalent length to the original CTS. However, there is a growing body of evidence indicating reliability and validity (Lucente, Fals-Stewart, Richards, & Goscha, 2001; Newton, Connelly, & Landsverk, 2001; Tuomi Jones, Ji, Beck, & Beck, 2002). On the other hand, there does not seem to be information about the reliability and validity of the CTS2 outside of the North American context in which they were developed. An instrument can have excellent psychometric properties in one sociocultural context and may not in another. The primary purpose of this article is to help answer the question of whether the demonstrated internal consistency reliability of the CTS2 in the North American context applies to other regions of the world. A secondary purpose, because the data on validity are more limited, is to present preliminary evidence on the cross-national construct validity of the CTS2.

THE INTERNATIONAL DATING VIOLENCE STUDY

The data for this article are from the International Dating Violence Study (IDVS). The IDVS is being conducted by members of a consortium of researchers at universities in every major world region except Africa. A detailed description of the study, including the questionnaire and all other key documents, is available on the Web site http://pubpages.unh.edu/~mas2. Some preliminary results have been published (Straus & Members of the International Dating Violence Research Consortium, 2004).

There is a core questionnaire that includes the CTS2, which each member of the IDVS research consortium translated. The members also agreed to back-translate the questionnaire to maintain conceptual equivalence (Straus, 1969) across the sites. However, there are no data on the percentage of those who actually back-translated. In addition, the consortium members added questions to measure variables that are uniquely important for their site or to measure constructs that are needed to test a theory of particular interest. These procedures allow the benefits of both standardized measures for all the sites and also the benefits of culturally informed investigations of unique issues at each university.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

The procedures followed to protect the rights and safety of the participants were reviewed by appropriate authorities at each university. These procedures explained the purpose of the study and the fact that the questionnaire contains questions on sensitive issues, including sexual relationships. To respect the privacy and the voluntary nature of participation, the instructions emphasized that respondents should turn in a blank questionnaire if they did not want to answer the questionnaire and that they were free to omit any question they did not wish to answer. The same information was printed on the cover page of the questionnaire.

METHOD

SAMPLES

The research was conducted at 33 university sites in 17 countries, as listed in Table 1. In accordance with the protocol on human subjects under which the study was conducted, only students who were 18 years or older were given the questionnaire. Because this article is about behavior toward a dating partner, only students who had been in a dating relationship lasting a month or more could be included. This varied from 100% to less than a third in Pune, India, where dating is not part of the culture. A total of 7,179 students met these criteria. The number of cases at each site ranged from 89 (Pune, India, where the percentage who met the criteria of having been in a dating relationship was low) to 550, with a mean of 218. However, respondents who omitted one or more of the questions needed for the CTS2 could not be included, and this reduced the sample size to about 6,700. The *n*s vary slightly from table to table and within tables for the reasons indicated in the table footnotes.

Table 1 gives the scores for each site on the four variables that were used as controls in partial correlation analyses. If this were entirely a North American study, race would have been included. However, a question on race could not be in the standard IDVS questionnaire because U.S. racial categories are not meaningful in most of the countries in this study.

The questionnaires were administered in classes taught by members of the consortium and in other classes for which they could make arrangements. Thus, it is a convenience sample. The students were told that participation was entirely voluntary and

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

TABLE 1
Characteristics of Students by University

University Site	n	*Percentage Female*	*Mean Age*	*Relationship Length (in months)*	*Mean Social Desirability Scale Score*
Total	7,179	71.4	22.0	13.7	34.0
Asia and Middle East					
HKG-HONGKONG	161	58.4	24.1	12.3	33.1
INDia-PUNE	89	70.8	22.7	13.1	32.9
ISRael-EMEKZYRL	346	82.7	23.1	12.5	34.3
KORea-PUSAN	237	60.3	24.6	10.5	32.0
SGP-SINGAPORE	223	70.0	25.1	17.3	32.9
Australia and New Zealand					
AUStralia-ADELAIDE	219	80.8	23.8	15.8	33.8
NZL Zealand-CHRISTCH	118	77.1	21.4	12.5	32.4
Europe					
BELgium-FLEMISH	448	77.2	20.4	14.6	34.0
CHE-Switzerland FRENCH	220	71.4	21.6	16.0	33.2
CHE-Switzerland GERMAN	135	77.0	19.4	14.2	35.0
DEU Germany-FREIBURG	165	57.6	23.7	13.4	32.0
GBR-SCOTLAND	218	84.4	21.9	13.9	33.7
NDL Netherlands-AMSTRDM	127	77.2	22.3	14.3	34.6
PRT-BRAGA	150	40.7	22.2	15.6	35.4
Latin America					
BRAzil-SAOPAULO	322	66.8	21.5	13.1	34.6
MEXico-JUAREZ	208	83.7	20.7	12.8	37.1
North America					
CANada-HAMILTON	245	86.5	21.5	15.2	33.5
CANada-LONDON	120	58.3	19.4	11.2	33.2
CANada-MONTREAL	292	78.8	23.7	17.1	34.6
CANada-TORONTO	218	67.0	20.4	12.8	34.2
CANada-WINNIPEG	141	89.4	22.1	15.3	33.2
USA-CINCINN	303	53.1	20.6	13.6	34.3
USA-WASHINGTON DC	84	84.5	20.5	14.3	33.2
USA-INDIANA	234	70.5	19.8	12.7	34.7
USA-LOUISIAN	128	67.2	21.4	12.8	36.3
USA-MISSISSP	221	90.0	28.4	18.7	35.5
USA-NH 1 (1998)	550	67.6	19.5	9.1	33.5
USA-NH 2 (2002)	293	74.1	20.8	13.7	34.6
USA-PENNSLVNA	215	75.8	20.0	11.3	33.8

(continued)

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

TABLE 1
(continued)

University Site	n	*Percentage Female*	*Mean Age*	*Relationship Length (in months)*	*Mean Social Desirability Scale Score*
USA-TX NCDCHS	109	72.5	20.8	12.7	33.1
USA-TX-MEXican American	242	62.0	24.7	16.3	35.5
USA-TX-Non MEXican	230	55.7	24.1	15.4	34.0
USA-UTAH	168	64.9	21.9	11.6	33.6

NOTE: The first three letters for each site are the abbreviations used by the United Nations. They are given in this table to serve as a key to the data points in Figures 1, 2, and 3. The two Swiss sites are for French and German speaking students. The two sites for USA-Texas are for Mexican Americans and Non–Mexican Americans. The two sites for USA-New Hampshire are for two different samples, 4 years apart.

that if they did not wish to participate, they should deposit the blank questionnaire in the same box provided for all questionnaires. Less than 1% chose this option. The results describe what was found for the students in those classes in each country and cannot be taken as representative of students in general. A specific example of the unrepresentative nature of the sample is that, as shown in Table 1, at most sites, about two thirds of the students are women. This is because the questionnaire was usually administered in psychology, sociology, and criminology classes, where women students predominate. None of the classes were focused on family violence. Some of the characteristics of the students in each site are given in Table 1.

DATA QUALITY CONTROL

The completed questionnaires were examined for questionable response patterns, such as reporting an injury but not reporting an assault as having occurred or cases with an implausible response, such as attacking partner with a knife or gun 10 or more times in the past year. About 4% of the cases were identified as questionable and were removed from the sample. Therefore, the results to be reported apply to CTS2 data that had first been filtered to delete questionable cases. This is not a limitation because filtering of this type should be a standard procedure.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

THE CTS2

The CTS2 include scales to measure physical assault, injury from assault by a partner, psychological aggression, sexual coercion, and negotiation. The theoretical basis of the instrument and the items in each scale are given in Straus et al. (1996). The CTS2 measure both perpetration by the respondent and victimization of the respondent. This article reports results for perpetration for all of the scales except injury, because the ability to measure perpetration is one of the most unique aspects of this instrument. Measuring perpetration may also be the most questionable aspect of the CTS2 because it requires respondents to disclose engaging in criminal behavior. However, for the injury scale, victimization of the respondent by the partner was used because the perpetrator may not know about the injuries inflicted, especially minor injuries, such as those in the item *felt physical pain that still hurt the next day because of a fight with my partner.*

SOCIAL DESIRABILITY SCALE

Research that uses self-reported data needs to take into account the tendency of some respondents to minimize socially undesirable behavior. This study used the Social Desirability Scale of the Personal and Relationships Profile (Straus, Hamby, Boney-McCoy, & Sugarman, 1999; Straus & Mouradian, 1999). This is a 13-item scale adapted from Reynolds short form of the Marlowe-Crowne Social Desirability Scale (Reynolds, 1982). The scale measures the degree to which a respondent tends to avoid disclosing undesirable behavior. The items in the scale consist of behaviors that are undesirable but true of almost everyone, such as "I sometimes feel resentful when I don't get my way." Consequently, the more of these almost universal items a respondent denies, the more likely the respondent is to also deny more seriously undesirable information, such as assaulting a partner and other forms of crime.

The response categories for each item are as follows: 1 = *strongly disagree*, 2 = *disagree*, 3 = *agree*, and 4 = *strongly agree*. The 13 items were summed. The higher the score, the greater the tendency to deny socially undesirable behavior. The theoretical range of the Social Desirability Scale is from 13 to 52. For this sample, the scores ranged from 18 to 52 (M = 34.2, SD = 4.8). The alpha coefficient of reliability for this sample is .70 (.67 for male students and

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

.71 for female students). Because differences between cultural groups in self-reports of violence against a partner might be a reflection of differences in willingness to report such behavior, partial correlation was used to control for score on the Social Desirability Scale.

OTHER MEASURES

Dominance. Dominance by one partner in a relationship was included as one of the variables to examine construct validity because a number of studies have found that dominance is an important risk factor for partner violence (So-Kum Tang, 1999; Straus, 1976, 1994). For this study, dominance was measured by the 9-item version of the Dominance Scale (Hamby, 1996) in the Personal and Relationships Profile (Straus et al., 1999; Straus & Mouradian, 1999). An example item is *my partner needs to remember that I am in charge*. The response categories are as follows: 1 = *strongly disagree*, 2 = *disagree*, 3 = *agree*, and 4 = *strongly agree*. The alpha coefficient of reliability for this sample is .55 (.58 for male students and .54 for female students).

Corporal punishment experienced as a child. Corporal punishment was included as one of the variables to examine construct validity because several studies have found that the more corporal punishment experienced as a child, the greater the probability later in life of physically assaulting a partner (Gershoff, 2002; Straus, 2001, 2005). To measure corporal punishment, the students were asked to rate the degree to which they agree with the following item: *When I was less than 12 years old, I was spanked or hit a lot by my mother or father.* The response categories were as follows: 1 = *strongly disagree*, 2 = *disagree*, 3 = *agree*, and 4 = *strongly agree*. Respondents who agreed or strongly agreed were coded as having experienced frequent corporal punishment.

Demographic characteristics. Because gender differences are so important for understanding violence between partners, all analyses were replicated for male and female students. The age of the respondent and the length of the relationship were included in the analyses as controls.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

SITES WITH VERY LOW OR 0 PREVALENCE OF VIOLENT BEHAVIORS

In the first exploratory analyses, an alpha coefficient of reliability of –.07 was found for the sexual coercion scale in Amsterdam. Inspection of the items revealed some items with a 0 prevalence rate and others with an extremely low prevalence rate (e.g., 3%). Items for which everyone has a score of 0 are dropped when alpha is computed, and items with extremely low prevalence rates are extremely skewed and therefore have a low correlation with other items. This combination drastically lowers the coefficient of reliability because the number of items in a scale and the size of the correlation between the items are the main determinants of the alpha coefficient. This explanation of the low alpha was tested by using the prevalence rate as an indicator of skewness. A low prevalence rate indicates extreme skewness. The partial correlation of the prevalence rate with the alpha for each scale, controlling for score on the Social Desirability Scale, was computed. The results revealed that for each of the three scales measuring rare events, the lower the prevalence rate, the lower the alpha reliability. (physical assault $r = .30$, injury $r = .45$, and sexual coercion $r = .45$).

The most extreme form of low prevalence is items with a 0 prevalence. This is again illustrated by Amsterdam, where four of the seven sexual coercion items have a prevalence of 0. Such items cannot be included in the scale, thus reducing the number of items in the scale to two. Because the second determinant of alpha is the number of items in a scale, this further reduces the alpha for sites with a low prevalence rate. This explanation was tested by computing the correlation between the number of 0 prevalence items and the alpha coefficient. The resulting correlations were physical assault $r = –.71$, injury $r = –.61$, and sexual coercion $r = –.85$.

These correlations support the idea that the extremely low alphas at a few sites occurred because the phenomena being measured were absent or nearly absent. Consequently, the tables in this article do not include coefficients for which alpha is inappropriate because there were two or more items with a 0 prevalence rate.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

TABLE 2
Alpha Coefficients of Reliability by Gender

Sample	n[a]	*Assault*	*Injury*	*Sex Coercion*	*Negotiation*	*Psychological Aggression*
Students						
All students	6,774	.88	.89	.82	.88	.74
Male	1,942	.93	.92	.84	.88	.78
Female	4,828	.86	.87	.78	.87	.72
Sites[b]						
M	33	.87	.87	.81	.85	.73
SD	33	.06	.09	.11	.07	.07
Range	33	.72 to .95	.57 to .98	.44 to .91	.63 to .97	.53 to .83

a. The *n*s for each scale vary slightly because of missing data on some questions.
b. The *n* in the sites section refers to the number of sites.

RESULTS

RELIABILITY

Pooled student data. The student section of Table 2 is based on pooling the students from all 33 universities who had no missing data and had been in a dating or cohabiting relationship in the past 12 months ($N = 7,179$). It shows remarkably high levels of reliability. Even the lowest coefficient (.74 for psychological aggression) exceeds the convention that sets .70 as indicating good reliability. The differences between the reliability of the CTS2 for male and female students are small, but each of them shows that the reliability is slightly higher for male students.

Mean of sites. The sites section of Table 2 gives the mean of the alpha coefficients for the 33 sites. The first row in that section gives the mean based on the total sample in each site. The mean alpha coefficients are all slightly lower than the alpha coefficients based on pooling all 6,744 students. This is because there are a few sites with low alpha coefficients.

Specific site alpha coefficients. Table 3 shows the alpha coefficients for each of the 33 university sites for all students and for males and females at each site.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

CONFOUNDING WITH SOCIAL DESIRABILITY RESPONSE BIAS

Although the CTS2 have high internal consistency reliability, that does not show that the scales are valid. For example, the high internal consistency might reflect a tendency of respondents to be consistent in avoiding disclosing violent behavior. Consequently, it is important to determine if the differences between sites are an artifact of confounding with willingness to disclose socially undesirable behavior and beliefs. This hypothesis was tested by computing the correlation of the Social Desirability Scale with each of the CTS2.

The right-hand column of Table 1 gives the mean Social Desirability Scale scores for each of the 33 university sites. Table 4 gives the correlation of the Social Desirability Scale with three of the CTS2 scales for the total sample and for each of the 33 university sites. These are the scales for which social desirability response set is the most serious threat to validity because the items ask about a criminal behavior. The correlations show that the higher the Social Desirability Scale score, the lower the score on the physical assault, injury, and sexual coercion scales, indicating that the Social Desirability Scale is operating as intended. However, the correlations are low. The mean correlation for the physical assault scale was –.17 (range = –.03 to –.23) and –.09 for injury (range = .00 to –.23). These correlations are not high enough to be an important threat to validity. Nevertheless, to be on the safe side, the correlations in the construct validity section controlled for score on the Social Desirability Scale.

CONSTRUCT VALIDITY

The previous sections have shown that in the 33 diverse settings in this study, the CTS2 have high reliability and are not importantly confounded with social desirability response bias. These are necessary characteristics, but they are not sufficient. There must also be evidence of validity. This section therefore gives examples of analyses that provide preliminary evidence on the construct validity of the assault and injury scales.

Evaluation of construct validity requires examining the correlation of the measure being evaluated with variables that are known to be related to the construct purportedly measured by the

(continued on page 423)

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

TABLE 3
Alpha Coefficients of Reliability for Specific Sites

Sample	n[a]	*Assault*	*Injury*	*Sex Coercion*	*Negotiation*	*Psychological Aggression*
Asia and Middle East						
HKG-HONGKONG, total	220	.91	.90	.87	.97	.82
Male	87	.95	.95	.93	.95	.79
Female	133	.85	—	—	.97	.84
IND-PUNE, total	74	.93	.92	.90	.89	.81
Male	18	.93	.97	.98	.92	.75
Female	57	.93	.91	.86	.88	.81
ISR-EMEKZYRL, total	318	.94	.94	.88	.87	.72
Male	59	.94	—	.70	.88	.72
Female	259	.94	.96	.92	.86	.72
KOR-PUSAN, total	194	.92	.94	.84	.82	.80
Male	77	.83	—	—	.84	.76
Female	117	.93	.96	.94	.81	.81
SGP-SINGAPORE, total	212	.83	—	.62	.86	.64
Male	66	.96	—	—	.87	.77
Female	146	—	—	—	.85	.55
Australia and New Zealand						
AUS-ADELAIDE, total	231	—	.57	—	.87	.73
Male	42	—	—	—	.91	.43
Female	189	—	.56	—	.86	.76
NZL-CHRISTCH, total	111	—	—	—	.82	.62
Male	24	—	—	—	—	—
Female	87	—	—	—	.84	.64

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

Europe						
BEL-FLEMISH, total	439	.79	.70	—	.88	.71
Male	99	.87	.78	—	.89	.70
Female	341	.71	—	—	.88	.71
CHE-FRENCH, total	190	.88	.98	.81	.93	.71
Male	49	.94	.99	.93	.97	.78
Female	140	—	—	—	.90	.66
CHE-GERMAN, total	122	—	—	—	.92	.64
Male	27	—	—	—	.91	.69
Female	96	—	—	—	.92	.64
DEU-FREIBURG, total	162	.80	.83	.85	.72	.68
Male	68	.80	—	.86	.74	.71
Female	94	—	.92	—	.70	.64
GBR-SCOTLAND, total	213	.72	—	.44	.83	.65
Male	32	—	—	—	.65	.65
Female	181	.72	—	.46	.84	.65
NDL-AMSTRDAM, total	134	—	—	—	.83	.53
Male	33	—	—	—	.88	.59
Female	100	—	—	—	.81	.49
PRT-BRAGA, total	139	—	—	—	.81	.70
Male	81	—	—	—	.77	—
Female	58	—	—	—	.87	.63
Latin America						
BRA-SAOPAULO, total	254	.91	.90	.81	.82	.76
Male	81	.94	.96	.86	.81	.80
Female	173	.89	.83	—	.82	.73
MEX-JUAREZ, total	215	.86	.89	.83	.91	.70
Male	35	.94	.95	.90	.86	.77
Female	180	.83	.89	.78	.92	.69

(continued)

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

TABLE 3 (continued)

Sample	n[a]	*Assault*	*Injury*	*Sex Coercion*	*Negotiation*	*Psychological Aggression*
North America						
CAN-HAMILTON, total	259	.90	.89	.81	.87	.74
Male	35	.98	1.00	.90	.89	.82
Female	224	.87	.86	.77	.87	.73
CAN-LONDON, total	131	.93	.93	.91	.84	.83
Male	56	.95	.93	.93	.78	.81
Female	75	.93	.93	.91	.87	.84
CAN-MONTREAL, total	285	.76	—	—	.87	.75
Male	59	—	—	—	.93	.83
Female	226	.78	—	—	.79	.72
CAN-TORONTO, total	230	.88	.87	.82	.89	.75
Male	76	.87	.79	.71	.88	.73
Female	154	.89	.89	.87	.89	.77
CAN-WINNIPEG, total	109	—	.81	—	.76	.68
Male	12	—	—	—	—	—
Female	98	—	—	—	.78	.68
USA-CINCINN, total	308	.95	.95	.91	.91	.83
Male	151	.97	.96	.93	.91	.88
Female	157	.85	.93	.81	.90	.74
USA-INDIANA, total	211	.92	.91	.86	.84	.80
Male	57	.95	.93	.91	.88	.85
Female	154	.89	.88	.83	.83	.77
USA-LOUISIAN, total	106	.93	.90	.85	.72	.76
Male	33	.97	.95	.91	.45	.85
Female	74	.90	.88	.78	.75	.70

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

USA-MISSISSP, total	226	.87	.88	.84	.83	.75
Male	22	.94	.92	.96	.69	.84
Female	204	.86	.87	.81	.84	.74
USA-NH 1, total	734	.82	.84	.69	.90	.70
Male	227	—	—	.55	.92	.67
Female	504	.81	.82	.67	.88	.69
USA-NH 2, total	276	.83	.82	.66	.86	.70
Male	63	.93	.93	.86	.74	.76
Female	213	—	.60	.35	.88	.67
USA-PENNSLVNA, total	179	.93	.92	.85	.89	.73
Male	40	.98	.91	.91	.86	.83
Female	139	.90	.93	.82	.90	.70
USA-TX NCDCHS, total	82	.90	.92	.87	.87	.76
Male	20	.94	.94	.89	—	.85
Female	63	—	.93	.88	.89	.72
USA-TX-MEX, total	225	.86	.84	.79	.88	.79
Male	86	.92	.90	.82	.89	.81
Female	139	—	—	—	.87	.78
USA-TX-N MEX, total	218	.92	.88	.82	.90	.80
Male	96	.93	.93	.85	.86	.84
Female	121	.90	.85	.79	.92	.77
USA-UTAH, total	179	—	—	—	.90	.74
Male	66	—	—	—	.90	.73
Female	113	—	—	—	.91	.76
USA-WASHING DC, total	79	.83	.92	.84	.63	.64
Male	13	—	—	—	.57	—
Female	66	.84	.94	.86	.68	.65

NOTE: — indicates that alpha could not be computed because of two or more items with a 0 prevalence.
a. *n*s vary slightly for each scale because of differences in missing data.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

TABLE 4
Correlation of CTS2 With Social Desirability Response Bias Scale

University Site	n[a]	*Physical Assault*	*Injury*	*Sexual Coercion*
Total	6,788	–.17**	–.09**	–.11**
Asia and Middle East				
HKG-HONGKONG	220	–.17**	–.15*	–.12*
IND-PUNE	73	–.07	–.03	–.07
ISR-EMEKZYRL	320	–.12*	–.12*	–.10*
KOR-PUSAN	194	–.26**	–.13*	.01
SGP-SINGAPORE	213	–.22**	–.06	–.09
Australia and New Zealand				
AUS-ADELAIDE	232	–.26**	–.15*	–.31**
NZL-CHRISTCH	111	–.24**	–.22*	–.15
Europe				
BEL-FLEMISH	441	–.21**	–.10*	–.10*
CHE-FRENCH	190	–.21**	.00	–.14*
CHE-GERMAN	123	–.08	.03	–.06
DEU-FREIBURG	161	–.04	–.13*	–.11
GBR-SCOTLAND	213	–.30**	–.15	–.18**
NDL-AMSTRDAM	135	–.21**	–.10	–.08
PRT-BRAGA	139	–.02	–.11	–.22**
Latin America				
BRA-SAOPAULO	257	–.19**	–.08	–.18**
MEX-JUAREZ	216	–.32**	–.22**	–.10
North America				
CAN-HAMILTON	259	–.19**	–.11*	–.13*
CAN-LONDON	130	–.12*	.04	–.01
CAN-MONTREAL	287	–.26**	–.14**	–.21**
CAN-TORONTO	231	–.17**	–.14*	.00
CAN-WINNIPEG	110	.03	–.16	–.14
USA-CINCINN	309	–.24**	–.16**	–.14**
USA-WASHINGTON DC	79	–.17	–.01	.07
USA-INDIANA	211	–.23**	–.20**	–.14*
USA-LOUISIAN	107	–.23**	–.16*	–.18*
USA-MISSISSP	226	–.23**	–.04	–.08
USA-NH 1	737	–.20**	–.07*	–.11**
USA-NH 2	278	–.13*	–.16**	–.07
USA-PENNSLVNA	179	–.09	–.06	–.04
USA-TX NCDCHS	82	–.14	–.09	–.17*
USA-TX-MEX	221	–.20**	–.23**	–.20**
USA-TX-N MEX	216	–.13*	–.03	–.09
USA-UTAH	180	–.30**	–.11	–.09

a. *n*s vary slightly for each scale because of differences in missing data.
*$p < .05$. **$p < .01$.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

instrument being evaluated or for which there are theoretical grounds for expecting it to be related (Campbell & Fiske, 1959). Correlations that fit the expected pattern contribute evidence of construct validity. Construct validity is a judgment based on the accumulation of correlations from numerous studies using the instrument being evaluated. For this reason, the analyses in this section are best viewed as examples of results that must continue to be found to conclude that the CTS2 have cross-cultural construct validity.

Because the IDVS is a macrolevel study in which the focus is on explaining differences between sites in partner violence, the construct validity analyses for this article used macrolevel data on the rates of violence in each site. The macrolevel variables were computed using the SPSS procedure AGGREGATE to create a file in which the cases are the 33 university sites. The variables in that file consist of the percentage of students at each site who physically assaulted a dating partner in the previous year, the percentage who were physically injured by a dating partner, the percentage who experienced corporal punishment by a parent, and the mean for the site on a measure of dominance by one partner in a dating relationship. The percentages for each site will also be referred to as rates. The results describe differences in the rates of partner violence between the 33 sites, not differences in violence between individual students.

Because of space limitation, only results involving the physical assault and injury scales are presented. These were chosen because they are the most widely used of the CTS2.

Correlation of assault and injury. The question of whether students at universities with high rates of students assaulting a dating partner also have high rates of injury inflicted by a dating partner is highly suited for examining construct validity, as defined above, because by definition, they are related. Consequently, if that correlation is not found, it would raise serious questions about the validity of either the assault scale or the injury scale. Figure 1 shows the predicted higher injury rates at universities with high assault rates. Because this might be a spurious correlation, partial correlation was used to control for differences between university sites in the variables listed in Table 1, which includes a control for scores on the social desirability scale that measures differences in willingness of respondents to disclose socially undesirable behavior. The correlations of .77 and .75 in the upper right corner of

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

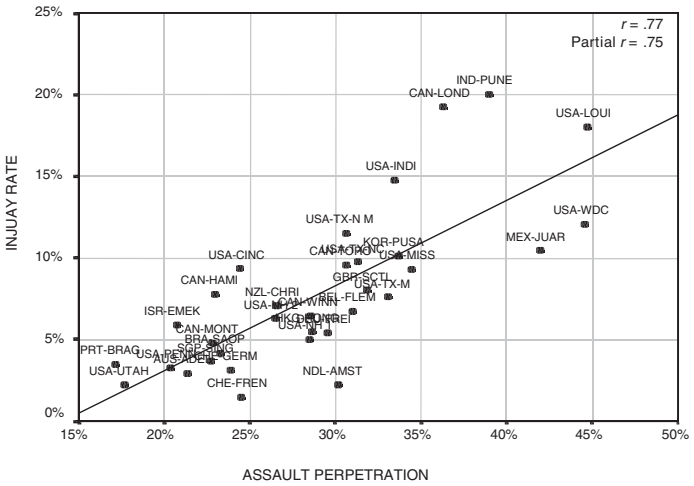

Figure 1: The Higher the Percentage of Those Who Assaulted a Partner, the More Partners Who Were Injured

NOTE: See Table 1 for key to site names.

Figure 1 show that the partial correlation between assault and injury was, as expected, lower than the zero order correlation, but only slightly lower. The use of partial correlation helps rule out the possibility that the differences between sites reflects site-to-site differences in willingness to disclose socially undesirable behavior rather than real differences in violence or differences in the gender composition of each site. In addition to showing that the measures of assault and injury are related in ways that must be present for the two scales to be valid, the high correlation with injury can be taken as evidence that the data on physical assault at the 33 sites refer to more than trivial events.

Correlation of corporal punishment with partner violence. The issue addressed by the second construct validity example is whether universities where a larger proportion of the students experienced corporal punishment as a child have higher partner assault rates than at universities where a smaller percentage experienced corporal punishment. The IDVS included a question that asked students whether they had been "spanked or hit a lot by" their parents when the respondent was younger than 12. At the

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

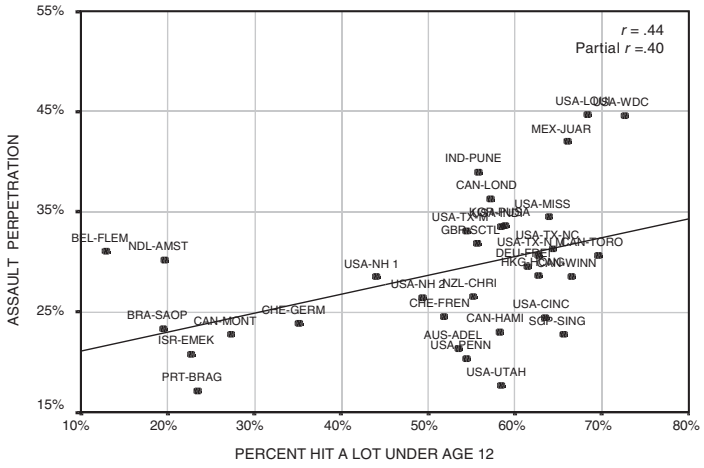

Figure 2: The More Corporal Punishment of Children, the Higher the Assault Rate

NOTE: See Table 1 for key to site names.

median university, 57% reported having been spanked or hit a lot as a child (range = 13% to 73%). Figure 2 shows that the larger the proportion of students who reported experiencing corporal punishment, the higher the percentage who had hit a dating partner in the past year. The correlations of .44 and .43 are much higher than the correlations typically found for the relation between childhood corporal punishment and violence as an adult, probably because these are macrolevel correlations that are usually higher than individual-level correlations. Except for the higher correlation, this result is consistent with many American studies, including prospective studies, which show that corporal punishment as a child is a risk factor for violence later in life (Gershoff, 2002; Straus, 2001, 2005) and therefore provides further data on the cross-cultural construct validity of the CTS2 Physical Assault Scale.

Dominance in dating relationships. A principle of conflict theory (Coser, 1967; Dahrendorf, 1959) is that inequality increases the risk of violence because the dominant person or group may use violence to maintain their position or the subordinate person or group may use violence to make the balance of power more equal. Feminist theory makes the same argument in respect to women who,

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

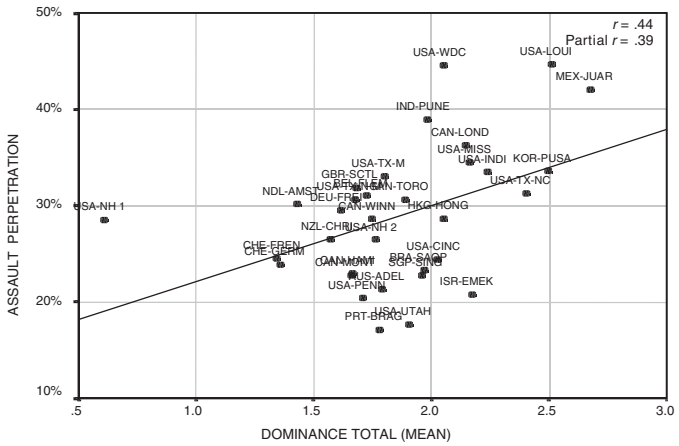

Figure 3: The More One Partner Is Dominant, the Higher the Assault Rate
NOTE: See Table 1 for key to site names.

the world over, tend to be subordinate in the family (Straus, 1976). Therefore, the more dating relationships are characterized by the dominance of one partner, the greater the probability of violence. The correlations of .44 and .39 in Figure 3 are consistent with this hypothesis and therefore provide an additional bit of evidence for the construct validity of the CTS2 Physical Assault Scale.

DISCUSSION

The results of this study of the dating relationships of students at 33 universities in 17 countries show that the alpha coefficients of reliability for the five CTS2 are generally high across all 33 universities, indicating that the CTS2 have cross-cultural reliability. The few instances where the reliability coefficients are low occurred because these are sites with a very low prevalence of partner violence. An extremely low prevalence rate is an extremely skewed distribution, and this reduces the size of the correlations. A low prevalence rate also results in items that cannot be included in the scale because they have a 0 prevalence rate. Because alpha is a function of the size of the correlations between items and the number of items in a scale, the combination of these two effects explains the instances of low alpha coefficients. More generally, the few

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

exceptions to the generally high alpha coefficients result from the inappropriateness of alpha as a measure of reliability when the items are extremely skewed rather than with cultural characteristics that make the data unreliable at those sites.

The correlations of the CTS2 with scores on a social desirability response bias scale were low, which is consistent with a meta-analysis of U.S. research on this issue (Sugarman & Hotaling, 1996). Consequently, it can be concluded that differences between national settings in the willingness of students to disclose violence is not an important threat to the cross-cultural validity of the CTS2, at least for studies of violence in the dating relationships of university students.

Evidence of construct validity was provided by scatter plots and partial correlations, which show that (a) universities with a high assault rate also tend to have a high injury rate; (b) the larger the percentage of students at a university who experienced frequent corporal punishment as a child, the higher the percentage of students who physically assaulted a dating partner; and (c) university sites where one partner tends to be dominant in dating relationships tend to have higher rates of assault on dating partners.

The large differences between sites in assault and injury rates shown in Figures 1, 2, and 3 and in a previous article (Straus & Members of the International Dating Violence Research Consortium, 2004) suggests that the CTS2 have adequate sensitivity to distinguish the level of violence against dating partners in different cultural contexts.

All analyses, where it was relevant, controlled for scores on a Social Desirability Scale and gender of respondent, thus making it unlikely that the results reflect university-to-university differences in willingness of students to disclose socially undesirable behavior or differences in the gender composition of each site.

LIMITATIONS

Although the results of this study indicate high alpha coefficient reliability across 33 diverse sites in 17 countries, alpha measures only internal consistency reliability. Temporal consistency as measured by test-retest data is arguably a more crucial aspect of reliability but was not measured. The absence of test-retest reliability is typical of social and psychological measures, including the CTS2. Of the more than 100 articles that have so far been published reporting results using this instrument, only 3 have reported data

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

on test-retest reliability (Straus, 2004), whereas more than 40 report alpha coefficients. Almost all those studies report reliabilities meeting or exceeding the conventional standard of an alpha of .70.

Concurrent validity, as indicated by the correlation of the CTS2 with other measures of the five constructs, was not investigated because priority was given to measuring more constructs needed to test theories of partner violence than to including additional measures of partner violence, especially given the belief that the CTS are the best available instruments (Archer, 1999; Grotevant & Carlson, 1989; Herzberger, 1991). Most social psychological measures also lack concurrent validity evidence for the same two reasons: that is, the absence of other validated measure of the same construct and because interview or testing time is usually available for only one measure of a construct. For the CTS2, as previously noted, there are now more than a hundred published studies, but only five examined concurrent validity (Straus, 2004). All five found that the CTS are correlated with other measures of approximately the same constructs.

The strong evidence of reliability and construct validity evidence found by this study does not necessarily mean that the CTS2 measure the same constructs in all sites. The question of cross-cultural conceptual invariance requires much more additional research. An approach that directly investigates the cross-cultural construct validity of the CTS2 is to replicate analyses that can be indicative of construct validity in each of the sites. This approach was followed in two analyses that were completed just as this article was in press. The first analyses tested the hypothesis that drinking problems are associated with violence toward a dating partner (Hines & Straus, 2004). The other analysis tested a hypothesis, based on Gottfredson and Hirschi's (1990) *A General Theory of Crime*, is that a low level of self-control is associated with a higher probability of all types of crime, including physically assaulting a dating partner. Both the drinking-problems study (Hines & Straus, 2004) and the self-control study (Rebellon & Straus, 2004) found the hypothesized relationship to partner violence as measured by the CTS in almost all the sites.

An important limitation is that the results refer to the behavior of university students and may not apply to the general population

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

and especially not to the low-income and low-education sectors of the population, where, at least in Euro-American societies, domestic violence rates are highest (Gelles & Straus, 1988; Straus, Gelles, & Steinmetz, 1980). Moreover, the results are also not representative of university students because they are based on a convenience sample consisting of the students who were enrolled in classes where the consortium member at each site was able to administer the questionnaire. Therefore, Figures 1 through 3 cannot be used to determine whether assault or injury rate of a specific country is lower or higher than other countries in the study. The value of these charts is in showing the relationship between variables, not in the absolute value of the variables. Because testing theoretically based hypotheses is the primary purpose of the cross-cultural use of the CTS, the results in this article provide evidence of validity, even though it is not possible to specify the population to which those results apply.

CONCLUSION

The alpha coefficients of reliability, the low degree of confounding with social desirability response sets, the construct validity examples in this article, and the construct validity evidence from two other studies all suggest that the CTS2 are appropriate instruments for measuring violence in partner relationships cross-culturally. However, all of this evidence refers to violence in the dating relationships of university students, whereas a definitive conclusion about reliability and validity requires consideration of a wide range of studies covering many applications of an instrument across a variety of populations. The CTS2 are becoming more widely used each year, and by mid 2004, more than 40 studies provided evidence of reliability (Straus, 2004), most of which are based on studies of nonstudent populations. Some of those studies are of different cultural groups within the United States, but none are cross-national. Nevertheless, the already extensive use of the CTS2 suggests that the needed wider range of evidence is likely to become available in relatively few years. In the meantime, the results presented in this article are sufficiently promising to encourage proceeding with that research.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

References

Archer, J. (1999). Assessment of the reliability of the Conflict Tactics Scales: A meta-analytic review. *Journal of Interpersonal Violence*, *14*(12), 1263-1289.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Coser, L. A. (1967). *Continuities in the study of social conflict*. New York: Free Press.

Dahrendorf, R. (1959). *Class and class conflict in industrial society*. Palo Alto, CA: Stanford University Press.

Gelles, R. J., & Straus, M. A. (1988). *Intimate violence*. New York: Simon & Schuster.

Gershoff, T. E. (2002). Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin*, *128*(4), 539-579.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Palo Alto, CA: Stanford University Press.

Grotevant, H. D., & Carlson, C. I. (1989). Conflict Tactics Scales (CTS). In *Family assessment: A guide to methods and measures* (pp. 203-299). New York: Guilford.

Hamby, S. L. (1996). The Dominance Scale: Preliminary psychometric properties. *Violence and Victims*, *11*(3), 199-212.

Herzberger, S. D. (1991). Conflict Tactics Scales (CTS). In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques, volume VII* (pp. 98-105). Austin, TX: Pro-ed.

Hines, D. A., & Straus, M. A. (2004). *Alcohol use and intimate partner violence: Cross-cultural associations and the moderating effect of antisocial personality*. Durham: Family Research Laboratory, University of New Hampshire.

Lucente, S. W., Fals-Stewart, W., Richards, H. J., & Goscha, J. (2001). Factor structure and reliability of the Revised Conflict Tactics Scales for incarcerated female substance abusers. *Journal of Family Violence*, *16*(4), 437-450.

Newton, R. R., Connelly, C. D., & Landsverk, J. A. (2001). An examination of measurement characteristics and factorial validity of the Revised Conflict Tactics Scale. *Educational and Psychological Measurement*, *61*(2), 317-335.

Rebellon, C. J., & Straus, M. A. (2004). *Self-control and violence against dating partners by university students in 16 countries*. Durham: Family Research Laboratory, University of New Hampshire.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, *38*(1), 119-125.

So-Kum Tang, C. (1999). Marital power and aggression in a community sample of Hong Kong Chinese families. *Journal of Interpersonal Violence*, *14*(6), 586.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

Straus, M. A. (1969). Phenomenal identity and conceptual equivalence of measurement in cross-national comparative research. *Journal of Marriage and the Family*, *31*, 233-239.
Straus, M. A. (1976). Sexual inequality, cultural norms, and wife-beating. In E. C. Viano (Ed.), *Victims and society* (pp. 543-559). Washington, DC: Visage.
Straus, M. A. (1990a). The Conflict Tactics Scales and its critics: An evaluation and new data on validity and reliability. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 49-73). New Brunswick, NJ: Transaction Publishing.
Straus, M. A. (1990b). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scale. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 29-47). New Brunswick, NJ: Transaction Publishing.
Straus, M. A. (1994). State-to-state differences in social inequality and social bonds in relation to assaults on wives in the United States. *Journal of Comparative Family Studies*, *25*(1), 7-24.
Straus, M. A. (2001). *Beating the devil out of them: Corporal punishment in American families and its effects on children* (2nd ed.). New Brunswick, NJ: Transaction Publishing.
Straus, M. A. (2004). *Bibliography and tabular summary of publications on the Revised Conflict Tactics Scales (CTS2 and CTSPC)*. Durham: Family Research Laboratory, University of New Hampshire.
Straus, M. A. (2005). *The primordial violence: Corporal punishment by parents, cognitive development, and crime*. Walnut Creek, CA: Alta Mira.
Straus, M. A., Gelles, R. J., & Steinmetz, S. K. (1980). *Behind closed doors: Violence in the American family*. New York: Anchor.
Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. (1999). *The personal and relationships profile (PRP)*. Durham: Family Research Laboratory, University of New Hampshire.
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., & Members of the International Dating Violence Research Consortium. (2004). Prevalence of violence against dating partners by university students worldwide. *Violence & Victims*, *10*(7), 790-811.
Straus, M. A., & Mouradian, V. E. (1999, November). *Preliminary psychometric data for the personal and relationships profile (PRP): A multiscale tool for clinical screening and research on partner violence*. Paper presented at the American Society of Criminology, Toronto, Ontario, Canada.
Sugarman, D. B., & Hotaling, G. T. (1996). Intimate violence and social desirability: A meta-analytic review. *Journal of Interpersonal Violence*, *12*(2), 275-290.
Tuomi Jones, N., 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.

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014

Yodanis, C., Hill, K., & Straus, M. A. (1997). *Tabular summary of methodological characteristics of research using the Conflict Tactics Scales.* Durham: Family Research Laboratory, University of New Hampshire.

Murray A. Straus is professor of sociology and founder and codirector of the family research laboratory at the University of New Hampshire. He has been president of the Society for the Study of Social Problems, Eastern Sociological Society, and National Council on Family Relations. He is the author or coauthor of more than 200 articles on family, research methods, and South Asia; and 15 books, including Stress, Culture, and Aggression *(Yale University Press, 1995),* Beating the Devil Out of Them: Corporal Punishment in American Families *(Transaction Publishing, 2001),* The Primordial Violence: Corporal Punishment, Cognitive Development, and Crime *(Alta Mira, 2005).*

Downloaded from ccr.sagepub.com at PORTLAND STATE UNIV on May 6, 2014