

Cross-cultural measurement equivalence of the Revised Conflict Tactics Scales (CTS2) Portuguese version used to identify violence within couples

Equivalência transcultural de mensuração da versão em português das *Revised Conflict Tactics Scales (CTS2)* na aferição de violência entre casais

Claudia Leite Moraes ¹
Michael E. Reichenheim ¹

¹ Departamento de Epidemiologia, Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. Rua São Francisco Xavier 524, Rio de Janeiro, RJ 20559-900, Brasil. clmoraes@ims.uerj.br michael@ims.uerj.br

Abstract *Following a previous evaluation of concept, item and semantic equivalences, this paper assesses the measurement equivalence between a Portuguese version of Revised Conflict Tactics Scales (CTS2) and the original instrument conceived in English. The CTS2 has been widely used to tap violence between couples. An intra-observer reliability evaluation involved 165 replications carried out within a 24-48 hour period. Kappa point-estimates were above 0.75 for all scales except sexual coercion. The analysis of internal consistency concerned 768 subjects with complete sets of items. Kuder-Richardson-20 estimates ranged from 0.65 to 0.86. Results were similar to those found in the original instrument in English for the negotiation, psychological aggression and physical violence scales, yet not so for the sexual coercion and injury scales. Factor analysis identified factors with a recognizable correspondence to the underlying dimensions, although a few inconsistencies were detected. For the assessment of construct validity (n = 528) associations between the instrument's scales were evaluated, as well as the relationships between violence and putative underlying dimensions. Overall, the findings suggest that the version can be used in the Brazilian context, although further investigation should be carried out to unveil some important remaining issues.*
Key words *Validity; Questionnaire; Spouse Abuse; Domestic Violence*

Resumo *Dando continuidade ao processo de adaptação transcultural do instrumento Revised Conflict Tactics Scales (CTS2), usado para a identificação da violência entre casais, o objetivo deste artigo é apreciar a equivalência de mensuração entre o original concebido em inglês e uma versão em português. A avaliação da confiabilidade intra-observador envolveu 165 replicações realizadas de 24 a 48 horas após a primeira abordagem. À exceção da escala de coerção sexual, todos as estimações de ponto do coeficiente kappa foram acima de 0,75. As estimativas do Kuder-Richardson-20 para a análise de consistência interna variaram de 0,65 a 0,86 (n = 768), sendo as relativas às escalas de negociação, agressão psicológica e violência física bem semelhantes às da versão original das CTS2, diferente das de coerção sexual e injúria. A análise fatorial (n = 768) identificou fatores com uma reconhecível correspondência com as dimensões subjacentes, ainda que algumas inconsistências tenham sido detectadas. Para estudar a validade de constructo (n = 528) foram avaliadas as associações entre as diversas escalas do instrumento e as relações entre violência e supostas dimensões subjacentes. Ainda que alguns pontos mereçam maior aprofundamento e mais investigação, os achados sugerem que a versão proposta pode ser usada no contexto brasileiro.*
Palavras-chave *Validade; Questionário; Maus-tratos Conjugais; Violência Doméstica*

Introduction

In different parts of the world, family violence has had an important impact on the health and mortality of children, adolescents, and women (Gelles, 1997; Heise et al., 1994; Kashani et al., 1992; Straus & Gelles, 1995). In Brazil, a clear picture has not yet been established, since information on the magnitude of family violence at the population level is still scarce. Still, even though available data are scant, they point to violence at the family level as an escalating problem (Heise, 1994; Soares, 1997).

The bio-psycho-social consequences of family violence have led to increasing interest within the scientific community in regard to prevention, follow-up of victims, and the development of pertinent research programs. Identification of violence in families is a major challenge. Thus, both the development of new and more accurate instruments and the improvement of existing ones for use at the clinical or research level have received worldwide attention, in addition to becoming the focus of several research groups (Aldarondo & Straus, 1994; Archer, 1999; Blinn-Pike & Mingus, 2000; Milner, 1994; Straus et al., 1998; Straus & Hamby, 1997; Weiss et al., 2000).

Accepting that in the field of family violence it is of great interest to compare profiles between cultures in different settings, it is essential that instruments conceived and extensively researched in one country also be thoroughly adapted for use in a new context. Local or regional nuances need to be contemplated in order to achieve cross-cultural equivalence.

This search for cross-cultural equivalence has motivated the Epidemiological Research Program on Family Violence at the State University of Rio de Janeiro to focus on studies aimed at providing Portuguese versions of several key instruments widely used in the field of family violence (Hasselmann & Reichenheim, 2001; Moraes et al., 2002; Reichenheim et al., 2000). Given the encouraging evaluations of the Conflict Tactics Scales (CTS) and their successful use in at least 20 countries (Archer, 1999; Assis, 1995; Hasselmann et al., 1998; Straus, 1979; Straus et al., 1996, 1998), the Revised Conflict Tactics Scales (CTS2) were selected to be adapted to the Portuguese language for use in Brazil. Straus et al. (1996) conceived the instrument in English in 1996 as an offshoot of the first CTS, now specifically directed to identify violence within married or dating couples and those in similar relationships. The CTS2 instrument consists of 78 items divided in five scales – negotiation, psychological aggression, physical

violence, sexual coercion, and injury – classified according to type and intensity. Items relate to the respondent and partner.

Adaptation of the CTS2 is based on guidelines proposed by Herdman et al. (1998). Accordingly, the process begins with a formal scrutiny of the relevance and equivalence of concepts and domains covered by the instrument as developed and used in the source population and that of the target culture where the new version is to be applied (concept equivalence). The model also prescribes an evaluation of the pertinence of each item comprising the original instrument in terms of its capacity to represent the underlying concepts and domains in the new context (item equivalence). Only then should the evaluation of linguistic equivalence be performed (semantic equivalence). The suitability of operational aspects in the new setting – form of administration, instructions, number of reply options, etc. – must also be observed (operational equivalence). In the sequence, the equivalence of psychometric properties between the original and the new version are formally evaluated (measurement equivalence). According to the authors, only after all these requisites have been met is it appropriate to state that the new version's cross-cultural adaptation has been achieved (functional equivalence).

In a previous paper, the first 3 phases of the process – concept, item, and semantic equivalence – were presented and a provisional Portuguese version proposed for use and further testing (Moraes et al., 2002). The present paper carries the adaptation process a step further, formally assessing the measurement equivalence between the original CTS2 and the new version. This article's main objective is to evaluate the Portuguese version's psychometric properties – reliability, internal consistency, and construct validity – and to compare the results with what has previously been reported on the original instrument in English.

Methods

Portuguese test version of the CTS2

As described in the previous section, the adaptation process of the Portuguese version of CTS2 has already been submitted to the first three steps of a formal evaluation. The assessment of conceptual and item equivalence included discussion groups with experts, as well as a comprehensive literature review to evaluate the pertinence of the concepts and items

covered by the original instrument in the Brazilian context. The ensuing appraisal of semantic equivalence consisted of two translations and respective back-translations; an independent assessment of the equivalence between the latter and the original instrument with regards to referential (literal) and general meaning; a discussion phase with specialists to re-examine and select items for the proposed version; and a pre-test on 774 women admitted to the hospital for childbirth. It was possible to ascertain conceptual, item, and semantic equivalence. In addition, acceptability of the Portuguese version was encouraging. Only three women refused to finish the interview. The proposed version is presented in Moraes et al. (2002) where the reader can also find the item numbers used throughout the rest of this paper. For the original instrument in English, refer to Straus et al. (1996). Note, however, that in this paper the numbering sequence is different since the respondents and partners receive consecutive numbers.

Fieldwork and data analysis

The present study is subsidiary to a hospital-based case-control study exploring the relationship between violence within families of pregnant women and premature childbirth. The fieldwork took place from March to September 2000 in three large public maternity wards in Rio de Janeiro. Specially trained interviewers collected data during the first 72 hours post-partum.

The analysis of intra-observer reliability relates to 165 subjects for whom replications were carried out by the same interviewer, 24-48 hours after the first meeting. Reliability is evaluated separately for each scale of the instrument, either using scores or as dichotomous variables. A positive event is defined as having at least one positive item in the scale, irrespective of whether the perpetrator is the woman or her partner. Two summary variables are also considered – Violence I (V1) and Violence II (V2). The first covers only items from the physical violence and injury scales. The second variable encompasses all items except those from the negotiation scale.

For dichotomized variables, reliability is estimated by the simple kappa statistic (Reliability I) for inter-rater agreement. Confidence intervals are calculated according to Fleiss et al. (1981) (equations 13.15 – 13.20) using an *ad hoc* Stata 7.0 routine (Reichenheim, 2001). For scores, kappas with quadratic weights are used (Reliability II) (Cohen, 1968). Confidence inter-

vals are calculated via bootstrap with 1,000 replications (Efron & Tibshirani, 1993; Stata-Corp, 2001). The option to use a weighted kappa is justified, since all scores are far from normal. The interpretation of kappa follows the guidelines by Shrout (1998), a development of the classification originally proposed by Landis & Koch (1977). Accordingly, agreement is classified as follows: virtually none ($\kappa < 0.1$); slight ($\kappa = 0.11-0.40$); fair ($\kappa = 0.41-0.60$); moderate ($\kappa = 0.61-0.80$); and substantial ($\kappa = 0.81-1.0$).

For analysis of internal consistency, data from all 768 subjects with complete sets of items is used. Each scale is evaluated separately for the respondent or partner and in aggregate form. All items are dealt with dichotomously. For the purpose of this analysis, injury items from the respondent and partner are swapped over since questions relate to consequences rather than actions. This exchange enables the assessment of internal consistency as seen from the perspective of the perpetrator, analyzing the deeds as well as any harmful effect brought about on the other. Calculations of point-estimates use the Kuder-Richardson-formula 20 – KR-20 (Streiner & Norman, 1995). Recognizing that KR-20 is always the lower limit of reliability (consistency) (Bleda & Tobias, 2000), a one-sided confidence interval is also calculated (Feldt, 1965). This means that there is a 95% chance that the reliability estimate will be higher than this computed value.

Factor analysis also relates to all 768 subjects with complete sets of items. The principal factors method is used with *varimax* rotation. Here too, items referring to the respondent are analyzed separately from partner. The results consider only factors with *eigenvalues* of and around 1, as suggested in Kleinbaum et al. (1988). All items are analyzed in binary form (Rummel, 1988).

Analysis of construct validity is confined to information from the main study's control group ($n = 528$). Restricting the subjects to controls prevents overestimating associations between violence and the other investigated constructs, since these are all putatively related to prematurity. Given the hospitals' broad coverage, one can assume that the sample is fairly representative of childbearing-age women using public health services in the city of Rio de Janeiro.

The assessment of construct validity contemplates two strategies. The first considers the relations between constructs represented by the scales. Five hypotheses involving the negotiation, psychological aggression, physical violence, sexual coercion and injury scales are

tested. Since these have been formerly explored by Straus et al. (1996), proximity to the results obtained in the original instrument is sought. Variables were selected that best represent the relationships at the theoretical level. The stated hypothesis and respective variables are shown in Table 4 in the results section.

The second strategy involves exploring the relations between the summary variables outlined above (V1 and V2) and other dimensions previously reported in the literature, namely, the woman's educational status and age; prenatal care; suspicion of alcohol and illicit drug abuse in the household; level of social support during pregnancy; and the family's socioeconomic status (CDC, 1994; Fagan & Browne, 1994; Gelles, 1997; Hillard, 1985; Stewart & Ccutti, 1993; Straus & Gelles, 1995; Webster & Palmer, 2000).

Variables corresponding to the underlying constructs are described in Table 5 found in the results section. Although most are self-evident, a few points need clarification: (a) The concept of educational status has been reduced to a four-level variable in accordance with the current Brazilian educational system. Thus, first level stands for 8 years of formal schooling and second level for an additional 3 years. The first category (incomplete first level) includes a few illiterates, but consists mostly of women with up to 3 or 4 years of schooling. The last category (complete second level or more) also includes those few who went on to graduate from university. (b) The cutoff point used for prenatal care follows the minimum acceptable number of consultations recommended by the Brazilian Ministry of Health (Brasil, 2000). (c) Suspicion of alcohol abuse is defined as either the respondent or her partner answering positively to two or more items from the CAGE (Cut-down; Annoyed; Guilty & Eye-opened) questionnaire (Mayfield et al., 1974). (d) Use of illicit drugs is determined by means of the NSDUQ (Non-Student Drug Use Questionnaire) instrument supported by the World Health Organization (Smart et al., 1981). A positive case is defined as either the respondent or the partner having consumed at least one listed illicit drug. The reference period for drug consumption is confined to pregnancy. (e) The level of social support is evaluated by the instrument developed by Sherbourne & Stewart (1991) and adapted for use in Brazil by Chor et al. (2001). Since there is no consensus as to how one should qualify social support, the overall score is first converted into quintiles. For ease of presentation, the three middle strata are then pooled together, given that the violence pattern be-

tween them is virtually indistinguishable. (f) An indicator describing environmental conditions for the family household is used as a proxy for the socioeconomic situation. The indicator is formed by four characteristics, namely, the household crowding level, expressed as the ratio between the number of dwellers and rooms; the predominant floor material; type of sewage disposal system; and the structure of domestic garbage disposal facilities.

Following the absence of normality mentioned previously, Kendall's τ_b is used to analyze the associations between ordinal variables (Kendall, 1970). The specific choice of τ_b is justified because the method appropriately corrects for ties (Agresti, 1984). Pearson's χ^2 test is used for categorical variables (Armitage & Berry, 1994). In the case of three-level variables, a score test for linear trend of the log odds is also employed in order to evaluate the presence of a violence gradient by strata (Clayton & Hills, 1995).

Data entry and its quality control were conducted, respectively, in Epi Info 6.04 (Dean et al., 1990) and Stata 7.0 (StataCorp, 2001). The latter was also used for data processing and analysis.

Results

Reliability

Intra-observer agreements are shown in Table 1. Both weighted and unweighted kappa point-estimates are around or above 0.75 for all scales except sexual coercion. Regarding sub-scales, one notes that in general the estimates are also high. The cognitive negotiation scale presents a much lower value when analyzed dichotomously.

Observing the contingency tables underlying the estimates (Table 1, column 3), one may notice that most disagreements relate to an assertion of the violent act or injury sustained as reported in the first interview and a denial in the second sitting. This occurs mainly in the minor sub-scales.

Internal consistency

Internal consistencies regarding the negotiation, psychological aggression, and physical violence scales are high and very similar to those found in Straus et al. (1996) (Table 2). Nevertheless, consistency estimates for the sexual coercion and injury scales are somewhat lower than in the original instrument.

By analyzing the items concerning respondent and partner separately, on the whole, one

Table 1

Intra-observer reliability of the Portuguese version of the Revised Conflict Tactics Scales (CTS2).

Scale	Sub-scale	2 nd		Reliability I* [Kappa (95% CI)]	Reliability II** [Kappa _w (95% CI)]
		1 st + +	+ -		
Negotiation	emotional	164	0	1.0 (-)	0.73 (0.61-0.84)
		0	1		
	cognitive	157	2	0.53 (0.17-0.89)	0.82 (0.72-0.88)
Psychological Aggression	minor	3	3	1.0 (-)	0.83 (0.60-0.82)
		164	0		
	both	0	1		
Psychological Aggression	severe	117	8	0.85 (0.77-0.94)	0.88 (0.82-0.92)
		1	39		
	both	74	5	0.86 (0.79-0.94)	0.90 (0.84-0.94)
Physical Violence	minor	6	80	0.85 (0.75-0.94)	0.90 (0.85-0.94)
		122	6		
	both	3	34		
Physical Violence	severe	53	15	0.74 (0.64-0.84)	0.87 (0.79-0.95)
		5	92		
	both	25	5	0.83 (0.72-0.94)	0.91 (0.82-0.96)
Sexual Coercion	minor	3	132	0.78 (0.69-0.88)	0.91 (0.84-0.96)
		57	13		
	both	4	91		
Sexual Coercion	severe	11	10	0.66 (0.46-0.85)	0.70 (0.38-0.86)
		0	144		
	both	2	2	0.80 (0.41-1.0)	0.80 (0-1.0)
Injury	minor	0	3	0.68 (0.50-0.86)	0.72 (0.46-0.88)
		12	10		
	both	0	143		
Injury	severe	17	8	0.74 (0.59-0.89)	0.73 (0.52-0.90)
		2	138		
	both	9	1	0.85 (0.68-1.0)	0.76 (0.50-0.94)
Violence I	minor	2	153	0.74 (0.60-0.89)	0.83 (0.65-0.93)
		20	8		
	both	3	134		
Violence I	severe	71	0	1.0 (-)	1.0 (-)
		0	94		
	both	33	0	1.0 (-)	1.0 (-)
Violence II	minor	0	132	1.0 (-)	1.0 (-)
		73	0		
	both	0	92		
Violence II	severe	124	5	0.90 (0.82-0.98)	0.90 (0.85-0.93)
		1	35		
	both	79	6	0.84 (0.76-0.92)	0.91 (0.87-0.96)
Violence II	severe	7	73	0.89 (0.80-0.98)	0.92 (0.88-0.95)
		128	3		
	both	3	31		

* Dichotomized variables; positives comprising a positive answer in at least one item of the respective sub-scale. Estimates using simple unweighted kappa.

** Scores comprising the sum of the scales' items. Estimates using quadratic weighted kappa.

Table 2

Internal consistency of the Portuguese version of the Revised Conflict Tactics Scales (CTS2) and the original instrument in English.

Scale (no. of items)	Woman perpetrator; male victim	Man perpetrator; female victim	Both	Straus et al.* (original)
Negotiation (6/12)	0.668** (0.637)***	0.719 (0.692)	0.807 (0.790)	0.86
Psychological aggression (8/16)	0.712 (0.686)	0.700 (0.672)	0.823 (0.807)	0.79
Physical violence (12/24)	0.779 (0.759)	0.860 (0.847)	0.864 (0.852)	0.86
Sexual coercion (7/14)	0.368 (0.310)	0.627 (0.592)	0.650 (0.619)	0.87
Injury (6/12)	0.387 (0.329)	0.681 (0.651)	0.659 (0.628)	0.95

* Original point-estimates using Cronbach's α (Straus et al., 1996).

** Kuder-Richardson-20 point-estimates (Streiner & Norman, 1995).

*** One-sided lower confidence interval (95%) (Feldt, 1965).

observes a slight decrease. Again, the sexual coercion and injury scales are exceptions. When the partner is the perpetrator, internal consistency is about twice as high as when it is the woman. It must be pointed out that in both situations, there is no crossing over between women's point-estimates and partners' lower 95% confidence limits.

When each item is systematically withdrawn from its respective scale, only item 8 stands out. Omitting this sexual coercion item raises the scale's point-estimate from 0.350 to 0.438, the only increase above 10% that is detected.

Factor analysis

Results of the factor analysis are presented in Table 3. Five factors are uncovered when the respondent (woman) is the perpetrator and four factors when the partner (man) is concerned. The total variance explained is, respectively, 97% and 82.9%.

The first factor relates, by and large, to physical violence, although several items pertaining to the psychological aggression scale also received moderate to high loadings in spite of the fact that this dimension is also portrayed by a separate factor. One item is particularly perceptible. Loadings for item 35 (*threatening to hit or throw something at the other*) are more in line with those relating to physical violence, yet quite diluted in regards to the original scale

(see factor 5 for respondents and factor 4 for partners).

Concentrating on the items originally allocated to the physical violence scale, only item 31 (*burning the other on purpose*) stands out with an almost null loading. Note that, while this item does not contribute to any other factor when women are perpetrators, a quite different picture emerges for the other group. For partners, item 31 seems to be very much in keeping with the second factor, which depicts sexual coercion, in spite of originally belonging to the physical violence scale. To a lesser degree, so does item 11 (*use of knife or gun*).

As mentioned, the second factor appears to represent sexual coercion. Apart from items 10 (*insisted on oral or anal sex by force*) and 24 (*insisted on sex by force*), which for respondents are non-informative due to a complete absence of positive answers, loadings tend to spread out among the other items of the scale. The exceptions are items 8 (*insisting on sex without a condom*) and 26 (*insisting on sex without use of force*).

The third factor clearly captures items encompassing the negotiation scale for both relations. The fourth, however, appears to be representing different dimensions. Injury-related items are favored when respondents are considered, whereas psychological aggression takes the fourth factor among partners. Although at first sight this seems to be describing a gender-

Table 3

Factor analysis of the Portuguese version of the Revised Conflict Tactics Scales (CTS2).

Scale	Item no.	Woman perpetrating					Man perpetrating			
		f1	f2	f3	f4	f5	f1	f2	f3	f4
Neg	1	0.026	0.016	0.481	-0.023	0.085	-0.124	-0.125	0.443	-0.181
	2	0.049	0.009	0.445	-0.032	0.012	-0.004	-0.046	0.494	-0.059
	7	-0.127	0.012	0.312	0.014	0.236	-0.207	-0.029	0.441	-0.314
	20	-0.035	0.013	0.629	-0.029	-0.066	-0.071	-0.060	0.620	0.023
	30	0.027	0.018	0.608	-0.015	-0.185	0.001	-0.051	0.601	0.080
	39	-0.039	0.024	0.499	0.010	0.059	-0.073	0.026	0.571	-0.105
Psy-M	3	0.320	0.016	0.003	-0.098	-0.540	0.291	0.065	-0.132	0.543
	18	0.288	0.018	0.136	-0.115	-0.560	0.208	0.044	0.009	0.563
	25	0.246	0.034	0.077	-0.144	-0.389	0.071	0.068	-0.032	0.503
	34	0.308	0.093	-0.003	-0.042	-0.362	0.323	0.038	-0.096	0.435
	35	0.632	0.081	0.009	-0.103	-0.152	0.556	0.218	0.031	0.231
Psy-S	13	0.149	0.041	0.050	-0.020	-0.455	0.097	0.088	-0.062	0.425
	15	0.371	-0.021	0.080	-0.034	-0.221	0.335	0.154	-0.072	0.203
	33	0.129	-0.003	-0.041	-0.006	-0.206	0.035	0.056	-0.053	0.250
Phy-M	4	0.600	-0.034	-0.032	0.017	-0.103	0.480	0.062	-0.087	0.104
	5	0.356	-0.012	0.015	0.091	-0.052	0.592	0.163	-0.089	0.133
	9	0.589	-0.036	-0.008	-0.055	-0.343	0.579	0.101	-0.015	0.351
	23	0.588	-0.017	0.030	-0.111	-0.174	0.520	0.096	0.016	0.371
	27	0.607	-0.024	0.048	-0.136	-0.124	0.741	0.116	-0.039	0.111
Phy-S	11	0.325	0.005	-0.010	-0.218	0.040	0.348	0.480	-0.176	-0.018
	14	0.643	-0.016	-0.008	-0.102	-0.009	0.770	0.025	-0.039	-0.016
	17	0.316	-0.002	-0.036	-0.080	-0.007	0.442	0.281	-0.127	-0.104
	19	0.502	0.160	-0.058	-0.077	-0.053	0.719	0.251	0.013	0.122
	22	0.420	0.006	-0.090	-0.067	0.099	0.752	0.090	-0.083	-0.042
	31	0.074	-0.009	-0.037	0.003	-0.027	0.129	0.693	-0.048	-0.063
	37	0.522	0.131	-0.041	-0.062	-0.039	0.635	0.161	-0.027	0.145
Sex-M	8	0.029	-0.007	-0.003	0.044	0.039	0.198	-0.001	-0.116	-0.023
	26	-0.001	0.270	0.109	-0.010	-0.223	0.192	0.160	0.014	0.246
	32	-0.024	0.612	0.052	0.008	-0.016	0.055	0.537	0.053	0.099
Sex-S	10	*	*	*	*	*	0.098	0.883	-0.056	-0.011
	24	*	*	*	*	*	0.280	0.615	-0.029	-0.021
	29	0.014	0.990	0.001	0.003	-0.006	0.082	0.812	0.030	0.122
	38	0.014	0.990	0.001	0.003	-0.006	0.153	0.565	0.011	0.161
Inj-M	6	0.270	-0.015	-0.081	-0.492	-0.072	0.346	0.037	0.125	0.257
	36	0.315	-0.010	0.000	-0.608	-0.095	0.335	-0.080	0.083	0.249
Inj-S	12	-0.027	-0.003	0.040	-0.549	-0.006	*	*	*	*
	16	0.066	-0.003	0.035	-0.608	-0.083	0.026	-0.032	0.015	0.218
	21	0.050	-0.009	0.061	-0.550	-0.059	-0.004	-0.014	0.052	0.048
	28	-0.001	-0.004	-0.001	-0.435	0.061	0.034	-0.036	-0.112	-0.003
% of variance explained by the factor		41.7	20.6	14.3	12.7	7.7	46.2	16.4	11.4	8.9

* Item not considered in the estimation due to lack of any positive answer.

Legend for scales: Neg = Negotiation; Psy-M = Minor psychological aggression; Psy-S = Severe psychological aggression; Phy-M = Minor physical violence; Phy-S = Severe physical violence; Sex-M = Minor sexual coercion; Sex-S = Severe sexual coercion; Inj-M = Minor injury and Inj-S = Severe injury.

dependent pattern, it is worth pointing out that a data-related artifact is at the core of this difference. Since there are virtually no injuries reported for partners and, as a result, a complete failure to represent this factor for the group, a re-location of psychological aggression items to the fourth factor automatically occurs, leaving the fifth factor virtually empty and void of any important contribution (hence not shown in Table 5). This point is further strengthened by the fact that, except for opposite signs, the respondent's fifth factor representing psychological aggression is quite similar to the relocated partner's factor 4. Besides, when artificially withdrawing the respondent's injury items from the data, an almost interchangeable respondent-partner pattern emerges. Now, a closer look at respondent's factor 5 and partner's factor 4 shows that, as with item 35 previously addressed, both items 15 (*destroying something belonging to the other*) and 33 (*accusing the other of being a lousy lover*) generate relatively small loadings. Whereas the former item may be regarded as also contributing to factor 1, the latter does not, nor to any other factor.

Construct validity

Results from the first evaluation strategy are shown in Table 4. The test hypotheses are described in the first column. The middle column contains the relationships actually tested. The next two columns present the correlations and respective p-values. Although point-estimates are sometimes low ($< \pm 0.10$), all tend in the theoretically expected direction and are statistically significant at the 10% level.

Associations between the two summary variables (V1 and V2) and selected construct-related variables are shown in Table 5. The overall picture is very consistent. Prevalence of violence is systematically higher in extreme strata of all variables inspected and significant correlations are in agreement with theoretically expected tendencies.

Discussion

Focusing the analysis on ordinal scales and disregarding the intensity of the events, reliability estimates are, by and large, quite substantial. When intensity is taken into account, intra-observer agreements tend to be slightly lower, oscillating between substantial and moderate. Even if estimates are again slightly lower, dichotomizing the scales still presents acceptable levels of reliability. The eccentric value

found in the cognitive negotiation sub-scale can be overlooked, since it is more a consequence of the idiosyncrasies in the estimator (κ) than a concrete finding. In fact, only 5 out of 165 cases are misclassified. All agreements are excellent when variables V1 and V2 are considered. It should not go unnoticed that summary variables like these are the most likely ones to be used in future substantive investigations.

Notice that for minor events, when disagreements occur, there is a regular pattern whereby the respondent tends to systematically deny the incident on replication. Three scenarios are plausible. In the first, asserting violence to begin with would only take place to please the interviewer, whereas the actual facts would only surface in the second interview. Rather more substantive than procedural, in the second scenario, an honest initial positive statement would be subsequently rejected following some overnight thoughts about the matter and a genuine distrust of the felt negative repercussions that disclosing any episode of violence would entail. In the third scenario, despite what was reported in the first meeting, the respondent would then consistently deny the event in order to speed up the replication interview (leading to what is known as *nay-saying bias*).

Although none of the three scenarios directly tells us anything about the Portuguese version's performance, this finding leads one to reflect on the CTS2's general properties and possibly those of other related instruments. More studies are needed to clarify the matter and to deal with the problem, should it be confirmed. One potentially helpful approach to deal with the third scenario described above would be to focus more effort on reducing the length of the instrument. This issue is discussed later in this article.

Returning to the examination of the adaptation process per se, its adequacy is supported by the fact that the pattern described above does not take place in relation to major events. None of the three scenarios applies here. On the contrary, the perception of violence as an emergency situation and the search for help may redouble the respondent's concentration, leading the person to disclose an incident as many times as deemed necessary. In contrast, minor violent acts are bound not to be noticed as problems, perhaps leading to carelessness with the accuracy of the information provided. The general pattern of systematically higher κ 's in the major sub-scales is consistent with this conjecture.

Table 4

Construct validity of the Portuguese version of the Revised Conflict Tactics Scales (CTS2) according to selected associations between scales and/or sub-scales.

Hypothesis	Tested association	τ_b^*	p-value
1) In couples where psychological aggression takes place, physical violence is more common	Severe psychological aggression vs. minor physical violence	0.405	0.000
	Severe psychological aggression vs. severe physical violence	0.326	0.000
2) In couples that resort to negotiation tactics to deal with their conflicts, there is a tendency for less severe physical violence or injuries to occur	Negotiation vs. severe physical violence	-0.095	0.009
	Negotiation vs. severe injury	-0.085	0.026
3) In couples where physical violence occurs, there is a tendency for sexual coercion to take place as well, especially when the man is the perpetrator	Physical violence of any type vs. sexual coercion of any type; man perpetrating	0.237	0.000
	Physical violence of any type vs. sexual coercion of any type; woman perpetrating	0.076	0.061
4) In couples where severe psychological aggression occurs, sexual coercion is more frequent, especially when the man is the perpetrator	Severe psychological aggression vs. sexual coercion of any type; man perpetrating	0.163	0.000
	Severe psychological aggression vs. sexual coercion of any type; woman perpetrating	0.099	0.015
5) Severe physical violence leads to severe injuries when the man is perpetrating but not when the woman is the agent	Severe physical violence vs. severe injury grave; man perpetrating	0.355	0.000
	Severe physical violence vs. severe injury; woman perpetrating	0.064	0.131

* Kendall's Tau-b point-estimates (Kendall, 1970).

Internal consistency estimates for the negotiation, psychological aggression, and physical violence scales are quite similar to those reported in the original CTS2 paper (Straus et al., 1996), as opposed to the sexual coercion and injury scales. Some explanations for those findings are in order. The first concerns the scarcity of information in the data. Focusing on the scale describing the injuries sustained by the partner ($KR-20 = 0.387$), for example, apart from the overall low prevalence of each event – only about 10% of the sample having at least one positive item –, positives are mostly concentrated in the two items related to minor events. In contrast, this pattern is rather different for the respondent's injury scale where the distribution of positive items is more spread out. Although some clustering in minor items remains, there are more positives in major injury items, as expected. Prevalence of severe injuries sustained by women is six times that of men (6.2% vs. 1.2%).

The small variability of injury scores for partners may also be responsible for the low $KR-20$ values. It should be pointed out that the virtual absence of severe injuries among partners is not surprising, since the consequences of a violent act in men are considerably less severe than in women, as depicted in Table 4 (5th

test hypothesis) and several previous studies (Barnett et al., 1997; Gelles, 1997; Hillard, 1985; Jasinski & Williams, 1998; Straus & Gelles, 1995; Webster & Palmer, 2000).

The same may be happening with the sexual coercion scale, and even more strikingly. For instance, the number of positive answers referring to partners is about twice as large if severity is ignored. Considering only the severe subscale, this ratio increases to ten. Small variability of scores can also be found here, since the majority of positives are concentrated in items 8 and 26, both pertaining to minor events. Another possibility that may partially explain the low internal consistency of the sexual coercion scale in particular is the apparent inappropriateness of item 8. Its withdrawal shoots the $KR-20$ estimate from 0.395 to 0.487, almost a 25% increase. This does not occur with the injury scale or, for that matter, any other scale. Internal consistency estimates remain roughly unchanged when one systematically withdraws each item. The third explanation is that the discrepancies found between the estimates of the two scales and the original are simply due to real differences between the adapted version and the CTS2 paper mentioned in Table 2. According to Straus et al. (1996), semantic modifications were introduced in several items of

Table 5

Construct validity of the Portuguese version of the Revised Conflict Tactics Scales (CTS2) according to associations between violence variables and selected construct variables.

Construct variable ^a	N	Violence I ^a		Violence II ^b	
		%	τ_b ^c	%	τ_b ^c
Educational status (woman)					
Incomplete first level	309	21.4		56.0	
Complete first level	86	17.4		53.5	
Incomplete second level	59	10.2		49.1	
Complete second level or more	75	10.7		37.3	
p-value ^d		0.056		0.033	
p-value ^e		0.007		0.005	
Correlation ^c			-0.089		-0.094
p-value			0.012		0.012
Adolescent (mother)					
Yes	165	24.2		57.6	
No	364	15.1		49.7	
p-value ^d		0.009		0.011	
Correlation ^c			-0.073		-0.050
p-value			0.036		0.121
Attendance, pre-natal care					
Less than six visits	254	22.0		59.8	
Six or more visits	257	14.4		45.1	
p-value ^d		0.029		0.001	
Correlation ^c			-0.083		-0.141
p-value			0.024		0.000
Suspicion of alcoholism (CAGE)					
Yes	123	29.3		71.5	
No	406	14.5		46.3	
p-value ^d		0.000		0.000	
Correlation ^c			0.153		0.186
p-value			0.000		0.000
Illicit drug consumption (NSDUQ)					
Yes	87	43.7		67.8	
No	442	12.9		49.1	
p-value ^d		0.000		0.001	
Correlation ^c			0.294		0.121
p-value			0.000		0.005
Social Support (MOS)					
Q1	112	28.6		61.6	
Q2 – Q4	311	17.0		54.6	
Q5	106	9.4		34.9	
p-value ^d		0.001		0.000	
p-value ^e		0.000		0.000	
Correlation ^c			-0.114		-0.136
p-value			0.001		0.000
Conditions of the household					
Bad	95	33.7		64.2	
Regular	202	16.3		52.0	
Good	226	13.3		47.3	
p-value ^d		0.000		0.023	
p-value ^e		0.000		0.008	
Correlation ^c			-0.128		-0.076
p-value			0.000		0.030

^a Summary variable covering items from the physical violence and injury scales.

^b Summary variable covering items from the psychological aggression, physical violence, sexual coercion, and injury scales.

^c Kendall's Tau-b point-estimates (Kendall, 1970).

^d Pearson's chi-square test (Armitage & Berry, 1994).

^e Test for trend for odds (Clayton & Hills, 1995).

the two scales after the reported psychometric results had been established. Thus, the estimates used for comparison do not strictly refer to the instrument's final edition to which the current adaptation applies.

Overall, the factor analysis identified factors with a recognizable correspondence to the underlying dimensions originally proposed by Straus et al. (1996). Nonetheless, it should be pointed out that, although quite able to capture the larger dimensions purported by the instrument, no distinction could be made in regards to sub-scales. Since, to the best of the authors' knowledge, no empirical evidence arising from a factor analysis of the original instrument in English is available, it is difficult to know whether such a discrimination should be expected but failed to materialize in the present study or, rather, if the differentiation between minor and severe events is to be made solely on the basis of theory, whatever the culture. This important point clearly needs further light. In any case, a comprehensible pattern can still be identified which is encouraging so far as the establishment of cross-cultural equivalence is concerned.

Comments on the peculiarities of some items and scales are also due. Confining the discussion to the results for the respondent's attitudes, a distinctive agglomeration of high loadings is observed for the physical violence items in the first factor and, in addition, there appears to be some influence by psychological aggression items. This is not an unexpected finding if one bears the related literature which points to high correlations between the two dimensions (Jasinski & Williams, 1998; Soares, 1997; Straus et al., 1996). The same concurrence has been detected in one of the confirmed hypothesis explored to assess construct validity. On the other hand, the existence of a separate factor representing psychological aggression should not go unnoticed. This means that both dimensions are effectively discriminated, once again underlining the suitability of the adaptation process.

The divergences detected for some items in relation to their respective scales deserve clarification. The similarity of the loadings for item 35 and those corresponding to the physical violence set could well be explained by the fact that this is the only item from the psychological aggression set bringing up a violent act that, taken a step further, may escalate and end up in corporal fighting. Threatening to hit or throw something at the partner (or vice versa) is quite different from the other items in the scale which refer strictly to offensive behavior

such as shouting, yelling, swearing, or “turning one’s back and walking away”. It is worth pointing out that the loading of item 35 is not much different from another that goes beyond intimidation and asks if something was actually thrown. Fifty-three percent of the respondents who made threats ended up throwing something, which clearly shows the synchronicity of the two practices. If this peculiarity is limited to the Portuguese version and the local culture or whether it also relates to the original CTS2 remains to be established. At any rate, item 35 does not appear to be causing much of a problem, as conveyed by the internal consistency analysis. Its withdrawal only reduces the α -20 estimate by 3.8% when the perpetrator is the respondent, and by 1.6% when it is the partner. These are negligible amounts by all means.

The two items relating to physical violence yet peculiarly contributing to factor 2 among partners also merit explanation. So far as item 31 is concerned not much can be inferred since there was only one positive subject who, perhaps incidentally, happened to also answer positively to four out of six sexual coercion items. Nevertheless, it should be pointed out that following a tentative substitution by another item in which the scope of the original item is narrowed down to specifics such as burning with a clothes iron, cigarette, or similar hot object, the odd loading of 0.693 is now reduced to 0.033, which may suggest there is some benefit from re-phrasing the question. The explanation for item 11 is rather more substantive. Apart from there being more information in the data, the majority of positive replies to threatening or using a gun to force sex, also used the latter in general as means to resolve conflicts. Testing hypothesis 3 presented in Table 4 has captured the same relationship. This shows that item 11, although valuable on its own right to factor 1 (physical violence), is redundant so far as the contribution to factor 2 is concerned.

As to the second factor that apparently captures sexual coercion, item 8 stands out. The item’s inadequacy, already mentioned in the discussion on internal consistency, again corroborates findings from previous phases of the adaptation process, especially with regard to the scrutiny of item equivalence and some important impressions that came up during the version’s pre-testing (Moraes et al., 2002). Since condom use is not widespread among couples in stable relationships, and given that this is the most common type of relationship found in the data, it is not surprising that interviewed women tend to simply dismiss the man’s insisting on sex without a condom as an act of coer-

cion. Possibly the same occurs with regard to item 26. In future versions, replacing the items with others more in tune with the Brazilian socio-cultural context should be considered.

Turning to the psychological aggression scale, the small loading for item 33 appears to confirm the suspicion raised in the previous phases of the adaptation process that, in the Brazilian context, accusing the other of being a lousy lover does not really provoke a confrontation, as first envisaged by the authors of the original CTS2. Even before this hard psychometric evidence, Moraes et al. (2002) identified the need to fine-tune or even replace this item altogether in future versions. Ratifying the propositions made in this previous study, in addition to a formal evaluation of measurement equivalence of any possible new substitution items, qualitative evaluations involving subjects from the population where the instrument is to be used should also be pursued.

Whereas item 33 is off mark whatever factor is concerned, item 15 – which also plays an unimportant role with regards to describing specifically psychological aggression – seems to offer some contribution to factor 1. At a closer look this should not be unexpected. Destroying belongings go well beyond verbal deeds, also involving negative physical acts. Perhaps this split contribution in terms of loadings is telling us exactly that. Further evaluations to shed some light on this matter may be encouraged.

Results obtained in the evaluation of construct validity also provide good evidence for the adequacy of the conversion of the CTS2 into Portuguese, since all associations scrutinized are consistent with both theory and other empirical studies. The related literature characterizes victims as being preferentially adolescent, poorly-educated females living in precarious socioeconomic conditions, involved in relationships where alcohol and illicit drug abuse are common, and where opportunities for contact with health services are limited due to low levels of prenatal care (CDC, 1994; Gielen et al., 1994; Heise et al., 1994; Hillard, 1985; O’Campo et al., 1995; Stewart & Cecutti, 1993; Webster et al., 1994; Webster & Palmer, 2000). Precisely the same picture was found in the present study.

Adequacy of the adaptation process is strengthened when some associations between the domains covered by the five scales are examined. Although the leniency of the significance levels and, at times, the weakness of some correlations must be acknowledged, on the whole, the postulated hypothesis are acceptably upheld. The positive associations between the severe psychological aggression sub-

scale and physical violence, whether minor or severe, corroborate the hypothesis that couples who are accustomed to initiating attempts at resolving conflicts by means of offensive psychological attacks tend to subsequently escalate to physical violence. The negative associations observed between the negotiation scale and the physical violence and/or severe injury scales support the hypothesis that there is less serious violence between couples that are used to resolving their misunderstandings and disagreements through demonstrations of appreciation and tolerance. The disparity between the consequences of physical aggression perpetrated by men and women, as depicted by a positive correlation between severe violence and injury when the man is the agent and a total lack of association when the woman is the perpetrator, also supports the adequacy of the adaptation process.

The present findings must also be examined in the light of some operational decisions that had to be made in the study. First, unlike the original CTS2, conceived as a self-reporting device, the current version was adjusted to the third person for face-to-face interviews to keep it in line with local practice. Since literacy is still not universal in Brazil, most epidemiological studies tend to be conducted using interviews. Second, in line with the requirements of the main case-control study, the recall reference period was narrowed down to cover only the months of pregnancy. Third, the analysis involves exclusively two-level items (happened/did not happen), again a practical constraint imposed by the fact that information on violence was collected along with other aspects as part of an extensive, time-consuming multi-dimensional questionnaire. Therefore, further studies using the originally proposed form, recall period, and scoring system ought to be carried out. In addition, it would be sensible to test the Portuguese version in other populations and contexts, since this study is restricted to women in their first postpartum hours and confined to public health services.

As hinted before, another interesting development would be to invest in a reduced edition of the instrument for use in health service-re-

lated evaluations or in multi-thematic investigations requiring multi-dimensional questionnaires. However, proposing a shortened version demands new trials, especially with regard to psychometric evaluations and entailing equivalence comparisons with both the original CTS2 and the complete version. It is worth recalling that all efforts spent in formally adapting an instrument conceived in another cultural context are fully rewarded by the external comparability that properly attuned and valid instruments provide. This is even more important in the field of family violence, since many gaps still need to be filled, particularly in countries like Brazil, where only now is the problem coming out into the open.

At present, the literature on cross-cultural adaptation methods in the field of family violence is still fuzzy. Even where some guidelines exist, there is no consensus as to the appropriate procedures required to adapt instruments to new contexts, cultures, or languages. Strategies ranging from simple translation by field researchers themselves to more detailed processes, like the one proposed by Herdman et al. (1998), are still under debate. Nonetheless, this model was chosen because it is comprehensive and unambiguously stresses the need to appreciate several nuances of the adaptation process. It should be born in mind that it is only one amongst several routines available in the literature, particularly in the field of Health-Related Quality of Life (Guillemin et al., 1993; Perreger et al., 1999).

Evaluation of the psychometric properties of the Portuguese version of the CTS2 complements the work initiated with the assessment of concept, item, and semantic equivalence. As a whole, the sensible results are encouraging and strengthen previous findings, indicating that the version can be recommended for use in research programs on family violence in the Brazilian context. Still, the current evaluation of measurement equivalence identified some important points that deserve discussion and further investigation in order to improve the proposed version and enhance the quality of research in the field at large.

Acknowledgments

C. L. M. has been generously supported throughout the project by a research grant from the Escola Nacional de Saúde Pública – ENSP (National School of Public Health)/Fundação Oswaldo Cruz – FIOCRUZ (Oswaldo Cruz Foundation). M. E. R. was partially funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Brazilian National Research Council), grant 300234/94-5. The authors are also thankful to Maria Helena Hasselmann who kindly shared her time and opinions and the field

staff (Alessandra Veggi, Fernanda Faria de Almeida, Isabella Mota Rezende, Joelma Castro and Juliana Toshica Kunisawa) for their outstanding work and commitment. Funding for this project was provided in part by a grant from CNPq, grant 663073/9987 (PRONEX Project) and in part by Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ (Rio de Janeiro State Research Foundation), grants E-26/171.223/98 and E-26/150.893/99.

References

- AGRESTI, A., 1984. *Analysis of Ordinal Categorical Data*. New York: John Wiley & Sons.
- ALDARONDO, E. & STRAUS, M. A., 1994. Screening for physical violence in couple therapy: Methodological, practical, and ethical considerations. *Family Process*, 33:425-439.
- ARCHER, J., 1999. Assessment of the reliability of the conflict tactics scales: A meta-analytic review. *Journal of Interpersonal Violence*, 14:1263-1289.
- ARMITAGE, P. & BERRY, G., 1994. *Statistical Methods in Medical Research*. 3rd Ed. London: Blackwell Scientific Publications.
- ASSIS, S. G., 1995. *Trajetória Sócio-Epidemiológica da Violência contra Crianças e Adolescentes: Metas de Prevenção e Promoção*. Tese de Doutorado, Rio de Janeiro: Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz.
- BARNETT, O. W.; MILLER-PERRIN, C. L. & PERRIN, R., 1997. *Family Violence across the Lifespan: An Introduction*. 1st Ed. London: Sage Publications.
- BLEDA, M.-J. & TOBIAS, A., 2000. Cronbach's alpha one-sided confidence interval (insert sg143). *Stata Technical Bulletin*, 56:26-27.
- BLINN-PIKE, L. & MINGUS, S., 2000. The internal consistency of the Child Abuse Potential Inventory with adolescent mothers. *Journal of Adolescence*, 23:107-111.
- CDC (Centers for Disease Control and Prevention), 1994. Physical violence during the 12 months preceding childbirth – Alaska, Maine, Oklahoma, and West Virginia, 1990-1991. *JAMA*, 271:1152-1153.
- CHOR, D.; GRIEP, R. H.; LOPES, C. S. & FAERSTEIN, E., 2001. Medidas de rede e apoio social no Estudo Pró-Saúde: Pré-testes e estudo piloto. *Cadernos de Saúde Pública*, 17:109-118.
- CLAYTON, D. G. & HILLS, M., 1995. *Statistical Models in Epidemiology*. Oxford: Oxford University Press.
- COHEN, J., 1968. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70:213-220.
- DEAN, A. G.; DEAN, J. A.; BURTON, A. H. & DICKER, R. C., 1990. *Epi Info, Version 6.04: A Word Processing, Database, and Statistics Program for Epidemiology on Micro-Computers*. Atlanta: Centers for Disease Control and Prevention.
- EFRON, B. & TIBSHIRANI, R., 1993. *An Introduction to the Bootstrap*. London: Chapman and Hall.
- FAGAN, J. & BROWNE, A., 1994. Violence between spouses and intimates: Physical aggression between women and men in intimate relationships. In: *Understanding and Preventing Violence* (A. J. Reiss Jr. & J. A. Roth, ed.), pp. 115-292. Washington, DC: National Academic Press.
- FELDT, L. J., 1965. The approximate sampling distribution of Kuder-Richardson reliability coefficient twenty. *Psychometrika*, 30:357-371.
- FLEISS, J. L., 1981. *Statistical Methods for Rates and Proportions*. 2nd Ed. New York: John Wiley & Sons.
- GELLES, R. J., 1997. *Intimate Violence in Families*. London: Sage Publications.
- GIELEN, A. C.; O'CAMPO, P.; FADEN, R. R.; KASS, N. & XUE, X., 1994. Interpersonal conflict and physical violence during the childbearing year. *Social Science and Medicine*, 39:781-787.
- GUILLEMIN, F.; BOMBARDIER, C. & BEATON, D., 1993. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology*, 46:1417-1432.
- HASSELMANN, M. H.; LOPES, C. S. & REICHENHEIM, M. E., 1998. Confiabilidade das aferições de um estudo sobre violência familiar e desnutrição severa na infância. *Revista de Saúde Pública*, 32:437-446.
- HASSELMANN, M. H. & REICHENHEIM, M. E., 2001. *Avaliação da Equivalência Semântica e de Mensuração da Versão em Português das "Conflict Tactics Scales – Form R" usada na Aferição de Violência Familiar*. Rio de Janeiro: Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. (mimeo.)
- HEISE, L., 1994. Gender-based abuse: The global epidemic. *Cadernos de Saúde Pública*, 10:135-145.
- HEISE, L. L.; PITANGUY, J. & GERMAIN, A., 1994. *Violence Against Women: The Hidden Health Burden*. Washington, DC: World Bank.
- HERDMAN, M.; FOX-RUSHBY, J. & BADIA, X., 1998. A model of equivalence in the cultural adaptation

- of HRQoL instruments: The universalist approach. *Quality of Life Research*, 7:323-335.
- HILLARD, P. J. A., 1985. Physical abuse in pregnancy. *Obstetrics & Gynecology*, 66:185-190.
- JASINSKI, J. L. & WILLIAMS, L. M., 1998. *Partner Violence: A Comprehensive Review of 20 Years of Research*. London: Sage Publications.
- KASHANI, J. H.; DANIE, A. E.; DANDOY, A. C. & HOLCOMB, W. R., 1992. Family violence: Impact on children. *Journal of American Academy of Child and Adolescence Psychiatry*, 31:181-189.
- KENDALL, M. G., 1970. *Rank Correlation Methods*. 4th Ed. London: Charles Griffin & Company.
- KLEINBAUM, D. G.; KUPPER, L. L. & MULLER, K. E., 1988. *Applied Regression Analysis and Other Multivariable Methods*. Boston: PWS-KENT Publishing Company.
- LANDIS, J. R. & KOCH, G. G., 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33:159-174.
- MAYFIELD, D.; McLEOD, G. & HALL, P., 1974. The CAGE questionnaire: Validation of a new alcoholism screening instrument. *American Journal of Psychiatry*, 131:1121-1123.
- MILNER, J. S., 1994. Assessing physical child abuse risk: The Child Abuse Potential Inventory. *Clinical Psychology Review*, 14:547-583.
- MORAES, C. L.; HASSELMANN, M. H. & REICHENHEIM, M. E., 2002. Adaptação transcultural para o português do instrumento "Revised Conflict Tactics Scales (CTS2)" utilizado para identificar a violência entre casais. *Cadernos de Saúde Pública*, 18:163-176.
- MS (Ministério da Saúde), 2000. *Programa Humanização do Parto: Humanização no Pré-Natal e Nascimento*. Brasília: MS.
- O'CAMPO, P.; GIELEN, A. C.; FADEN, R. R.; XUE, X.; KASS, N. & WANG, M.-C., 1995. Violence by male partners against women during the childbearing year: A contextual analysis. *American Journal of Public Health*, 85:1092-1097.
- PERNEGER, T. V.; LEPLÈGE, A. & ETTER, J.-F., 1999. Cross-cultural adaptation of a psychometric instrument: Two methods compared. *Journal of Clinical Epidemiology*, 52:1037-1046.
- REICHENHEIM, M. E., 2001. Confidence interval for the kappa-statistic of interrater agreement (kapci2.ado for Stata 7.0). Rio de Janeiro: Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. (available from author at michael@ims.uerj.br).
- REICHENHEIM, M. E.; MORAES, C. L. & HASSELMANN, M. H., 2000. Equivalência semântica da versão em português do instrumento *Abuse Assessment Screen* para rastrear a violência contra a mulher grávida. *Revista de Saúde Pública*, 34:610-616.
- RUMMEL, R. J., 1988. *Applied Factor Analysis*. 4th Ed. Evanston: Northwest University Press.
- SHERBOURNE, C. D. & STEWART, A. L., 1991. The MOS social support survey. *Social Science and Medicine*, 32:705-714.
- SHROUT, P. E., 1998. Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research*, 7:301-317.
- SMART, R. G.; ARIF, A.; HUGHES, P.; MEDINA-MORA, M. E.; NAVARATNAM, V.; VARMA, V. K. & WADUD, K. A., 1981. *Drugs Use among Non-Student Youth*. Geneva: World Health Organization.
- SOARES, B. M., 1997. *Mulheres Invisíveis: Violência Familiar e Formações Subjetivas*. Tese de Doutorado, Rio de Janeiro: Instituto Universitário de Pesquisas do Rio de Janeiro.
- STATA CORP, 2001. *Stata Statistical Software: Release 7.0*. College Station: Stata Corporation.
- STEWART, D. & CECUTTI, A., 1993. Physical abuse in pregnancy. *Canadian Medical Association Journal*, 149:1257-1263.
- STRAUS, M. A., 1979. Measuring intra-familial conflict and violence: The conflict tactics (CT) scales. *Journal of Marriage and Family*, 41:75-88.
- STRAUS, M. A. & GELLES, R. J., 1995. *Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8145 Families*. New Brunswick: Transaction Publisher.
- STRAUS, M. A. & HAMBY, S. L., 1997. Measuring physical and psychological maltreatment of children with the Conflict Tactics Scales. In: *Out of Darkness: Contemporary Perspectives on Family Violence* (G. K. Kantor & J. L. Jasinski, ed.), pp. 119-135, Thousand Oaks: Sage Publications.
- STRAUS, M. A.; HAMBY, S. L.; BONEY-MCCOY, S. & SUGARMAN, D. B., 1996. The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*, 17:283-316.
- STRAUS, M. A.; HAMBY, S. H.; FINKELHOR, D.; MOORE, D. W. & RUNYAN, D., 1998. Identification of child maltreatment with parent-child conflict tactics scales: Development and psychometric data for a national sample of American parents. *Child Abuse & Neglect*, 22:249-270.
- STREINER, D. L. & NORMAN, G. R., 1995. *Health Measurement Scales. A Practical Guide to Their Development and Use*. 2nd Ed. Oxford: Oxford University Press.
- WEBSTER, J.; SWEETT, H. & STOLZ, T. A., 1994. Domestic violence in pregnancy: A prevalence study. *Medical Journal of Australia*, 161:466-477.
- WEBSTER, J. J. & PALMER, R. L., 2000. The childhood and family background of women with clinical eating disorders: A comparison with women with major depression and women without psychiatric disorder. *Psychological Medicine*, 30:53-60.
- WEISS, S.; GARZA, A.; CASALETTO, J.; STRATTON, M.; ERNST, A.; BLANTON, D. & NICK, T. G., 2000. The out-of-hospital use of a domestic violence screen for assessing patient risk. *Prehospital Emergency Care*, 4:24-7.

Submitted on 22 June 2001

Final version resubmitted on 26 November 2001

Approved on 15 January 2002