Toward a Better Estimate of the Prevalence of Partner Abuse: Adjusting Rates Based on the Sensitivity of the Conflict Tactics Scale

Richard E. Heyman and Karin A. Schlee State University of New York at Stony Brook

Studies of spousal aggression, such as the national studies typically cited for prevalence rates (M. A. Straus & R. J. Gelles, 1986; M. A. Straus, R. J. Gelles, & S. K. Steinmetz, 1980), frequently use reports from only one spouse to calculate yearly prevalence. To date, no correction factor exists to help one estimate what the rates would have been had reports from both spouses been available. In this study, the authors calculate the epidemiological sensitivity of the Conflict Tactics Scale (M. A. Straus, 1979) in clinical and newly married samples and use the sensitivity figures to provide a correction equation. Correction factors are also provided for three previously published studies of interspousal agreement. The equations provided can be used to make rough estimates of the rate of male-to-female aggression when data from only one spouse are available.

Clinicians and researchers obtaining reports of woman abuse are frequently confronted with a problem: When information is available from only one partner, what is the best estimate of the true rate¹ of woman abuse? Indeed, the large sociological studies on the rates of woman abuse (e.g., Straus & Gelles, 1986; Straus, Gelles, & Steinmetz, 1980) relied on reports from one partner only. Clinicians working with only one partner do not have a heuristic for guessing how much aggression would, on average, be identified if the other partner were queried.

Because partner abuse is a socially undesirable behavior, it is reasonable to assume that data obtained from one spouse will provide the lower bound for the true rate of abuse. Despite a handful of articles on the interspousal reliability of abuse reports, no article to date has provided a simple but important piece of information:

What is the formula for a more reasonable estimate of the true rate of partner abuse, given reports by only one spouse? This formula can be easily derived by calculating the epidemiological sensitivity of the measure, and it can be used to estimate the number of uncaught cases when assessing sole respondents.

Previous studies on interspousal reliability have illustrated that marital partners do not concur on reports of physical aggression (Jouriles & O'Leary, 1985; Szinovacz, 1983) and that agreement between spouses on occurrence is generally low to moderate (O'Brien, John, Margolin, & Erel, 1994; O'Leary & Arias, 1988; O'Leary, Vivian, & Malone, 1992). Overall, women report more violence than men for both themselves and their partners (Browning & Dutton, 1986; Jouriles & O'Leary, 1985; Lawrence, Heyman, & O'Leary, 1995; O'Leary et al., 1989; Straus & Gelles, 1990). In terms of agreement, women and men appear to be more concordant when reporting on mild to moderate

Preparation of this article was supported in part by National Institute of Mental Health Grant MH42488. We thank Kenneth Leonard, K. Daniel O'Leary, and Dina Vivian for the use of their data sets.

Correspondence concerning this article should be sent to Richard E. Heyman or Karin A. Schlee, University Marital Clinic, Department of Psychology, State University of New York, Stony Brook, New York 11794-2500. Electronic mail may be sent via the Internet to rheyman@psychl.psy.sunysb.edu.

¹ The phrase *true rate* is used throughout this article in the strict sense of psychometric true score, that is, the population value that measurements only approximate. Any measured rate includes the true rate plus measurement error. Using the phrase *true rate* does not imply that the Conflict Tactics Scale, or any other measure, is infallible or without problems.

forms of aggression. With severe aggression, however, women tend to report higher frequencies of male violence than men do (cf. Straus & Gelles, 1990). This finding is consistent with studies that demonstrated that women are more seriously affected by male aggression, both physically and psychologically, than vice versa (e.g., Cascardi, Langhinrichsen, & Vivian, 1992).

In sum, the agreement literature indicates that there is a significant discrepancy between spouses' reports of physical aggression that does not seem to be completely accounted for by gender of respondent or severity level. Thus, although researchers know that spouses do not agree, it is not yet known how to compensate for their disagreement. Unfortunately, most studies on relationship aggression report prevalence rates that are based on a single spouse's report. without introducing any adjustment for underreporting. A correction equation would (a) allow estimation of what the prevalence rates of well-known studies (e.g., Straus & Gelles, 1990) would have been had researchers interviewed both spouses and (b) provide a heuristic by which reports from only one spouse can be used to generate a better estimate of the true (i.e., couple-based) prevalence rate for the sample. Although such a correction factor would not provide a definitive number by which all researchers should multiply all single-respondent reports of partner aggression, it would provide a reasonable best guess as to how much aggression would have been reported.

Method

Participants

Sample 1: Clinic Sample

Couples presenting for marital therapy at the University Marital Therapy Clinic at the State University of New York at Stony Brook (n=256) comprised the clinical sample. Couples completed an assessment battery before beginning marital therapy. Demographics are provided in Table 1.

Community Early Married Sample (Stony Brook Community Sample)

Subsample 2a. This subsample participated in the first wave of a longitudinal study previously described in O'Leary et al. (1989). In brief, couples (n=393) planning marriage in Onondaga and Suffolk counties, New York, were recruited through major newspaper and radio announcements to participate in a longitudinal study of first marriages. Couples completed an assessment battery approximately 1 month before marriage and were paid for each wave of participation. Demographics are provided in Table 1.

Subsample 2b. Couples (n = 128) responded to newspaper advertisements seeking participants married 1 to 7 years for a study of marriage. We administered, by telephone, the Marital Adjustment Test and the Conflict Tactics Scale to assess for study eligibility. Men and women were screened separately to assure the confidentiality of their reports (i.e., so that their spouse would not know their responses). Demographics are provided in Table 1. Subsamples

Table 1
Demographics of Clinic Sample and Stony Brook Community Sample

Variable			Stony Brook Community Sample					
	Clinic sample		Samp	ole 2a	Sample 2b			
	M	SD	M	SD	M	SD		
N couples	256		393		128			
% married	8	36	100		100			
Men								
Age	37.7	10.30	25.2	3.67	34.5	8.44		
Education	13.7	2.69	14.6	2.34	13.5	2.85		
MAT	74.5	10.70	121.3	17.81	108.2	28.64		
Women						_0.0.		
Age	34.9	9.75	23.4	2.98	32.1	6.70		
Education	13.7	2.31	14.5	1.97	13.5	2.83		
MAT	72.0	14.20	124.0	17.34	108.8	26.21		
Family income	\$46,471°	\$27,770	\$30,565b	\$13,118	\$46,652ª	\$21,415		

Note. MAT = Marital Adjustment Test. aIn 1990–1994 dollars. bIn 1982 dollars.

2a and 2b were combined into a total Long Island early married sample (Sample 2).

Buffalo Newlywed Study (Leonard & Senchak, 1994; McLaughlin, Leonard, & Senchak, 1992). This study collected a generalizable sample of couples (n = 839) applying for marriage licenses in Buffalo, New York. Details about participant recruitment are described in detail in McLaughlin et al. (1992). Ninety percent of couples approached for screening at City Hall agreed to participate; complete data from both men and women were collected on 73% of those who agreed to participate. All participants were required to (a) speak English, (b) be marrying for the first time, and (c) have a male between the ages of 18 and 29 years. The sample was demographically representative of couples marrying for the first time in this Northeastern U.S. urban area (i.e., average age 23-24 years, 26% ethnic minorities, high school educated, employed).

Re-analysis 1, 2, & 3

To offer a comparison to the results of our samples, we later provide similar sensitivity calculations for three previously reported reliability studies: Bohannon, Dosser, and Lindley (1995), Szinovacz (1983), and Margolin (1987). Bohannon et al. (1995) reported on 94 volunteer couples from a military base. The average age of men was 29.6 years, and for women it was 28.8. Also, 19% were officers, 36% were noncommissioned officers, and 46% were enlisted personnel.

Szinovacz (1983) reported on a sample of community couples. The majority of the participants were White, middle-class couples in their first marriage. Sixty percent of couples were under 40 years old, and 40% were between 40 and 70 years old. Fifty-eight percent of the men had attended at least some college; however, no information on women's education was available.

Margolin (1987) reported on 103 couples. No other demographic data were published on this sample.

Measures

Marital Adjustment Test (MAT)

The MAT (Locke & Wallace, 1959) is one of the most widely used measures of marital adjustment. Nine of the 15 scale items assess the degree of disagreement on major marital issues. It is sensitive to changes in marital therapy (cf. O'Leary, 1987), and its convergent validity has been repeatedly demonstrated (e.g., Navran, 1967; Spanier, 1976). Scores can range from 2 to 158, with higher scores indicating higher levels of adjustment; a score of 100 has been the traditional cutoff point for marital distress. MAT scores are provided as a descriptive measure for the

clinic sample and the Stony Brook Community sample.

Conflict Tactics Scale (CTS)

The CTS (Straus, 1979) is an 18-item self-report inventory assessing the frequency (on a 0-6 scale) of a variety of functional (e.g., calmly discussing a problem), verbally aggressive (e.g., insults or swearing), and physically aggressive (e.g., hitting) conflict tactics. Reliability of the CTS is good (Cronbach's $\alpha = .87$ for men, .88 for women; Straus, 1979). It has been used in national surveys of the prevalence of marital aggression (Straus & Gelles, 1986; Straus et al., 1980). Factor analyses of the CTS (e.g., Barling, O'Leary, Jouriles, Vivian, & MacEwen, 1987) demonstrated that the measure comprises physical and psychological aggression factors. We used the standard (Straus, 1979) classification scheme for mild aggression (throw something; push, grab, shove; slap) and severe aggression (kick, bite, hit with fist; beat up; threaten with a gun or knife; use a gun or knife). If the man were reported to have engaged in any of the severe aggression items in the last year, he was classified as severely aggressive. If he were reported to have engaged in no severe aggression but at least one instance of mild aggression, he was classified as mildly aggressive. If there were no reports of aggression in the last year, he was classified as nonaggressive. Men's reports of male-to-female aggression were compared with women's reports of male-to-female aggression.

Although the CTS is the most widely used measure of partner aggression, it is not without its critics. In particular, the CTS does not measure patterns of violence, their context, or the psychological impact and physical injury resulting from acts of aggression. The interested reader should consult the original critiques (e.g., Dobash & Dobash, 1984; Kurz, 1993; Saunders, 1988) and responses (e.g., Straus, 1990, 1993).

Method of Calculating Sensitivity

Epidemiological sensitivity is the ratio of the cases detected by a measure divided by all true cases (i.e., true positives + false negatives) and can be calculated by the following formula:

sensitivity = true positives ÷ all true cases.

Table 2 displays the decision matrix. A traditional epidemiological decision matrix compares a test versus a gold standard for knowing who has a disease. In our case, however, there is no gold standard for knowing who truly is aggressive (short of videotaping couples 24 hours a day). The typical assumption in

Table 2
Decision Matrix for Calculating Sensitivity

, -	Spouse report			
Respondent report	1	2		
1. Aggressive	True positive	True positive ^a		
2. No aggression during last year	False negative	True negative		

^aTraditional epidemiological matrices would consider these cases as false positives.

partner abuse research is to consider all positive reports as accurate, regardless of the reporter.² Thus, there are no possible false positives (e.g., a man who reports aggression when his partner does not is still considered a true positive).

Once sensitivity is known, correction equations can be created by modifying the standard sensitivity equation to solve for all true cases. Thus, if sensitivity of men's reports = .75, we can estimate the couple-based level of male-to-female aggression from men's reports as follows:

all true cases = men's reports $\div 0.75$

or, to make the equation more salient,

all true cases = men's reports \times 1.33.

As shown in Table 2, one needs a 2×2 table to calculate sensitivity. We created two indexes: (a) spouses' reports of any male-to-female aggression and (b) spouses' reports of severe male-to-female aggression. We collapsed the original 3×3 (None \times Mild \times Severe) tables into two sets of 2×2 sensitivity tables: (a) no report of male-to-female aggression during the past year versus any (i.e., mild or severe) report of aggression and (b) no report of severe (i.e., none or mild-only) male-to-female aggression during the past year versus report of severe aggression.

Results and Discussion

Sensitivity statistics are presented in Tables 3 and 4 along with the percentage of reports that are by the man only, by the woman only, and by both partners. The last two columns in both tables present correction factors (i.e., $1 \div$ sensitivity). They offer separate corrections for reports of general occurrence of aggression (i.e., presence or absence of any aggression) and for reports specifically of severe aggression.

Correction factors are provided as a heuristic to guess what the rate of aggression would have been had reports from both partners been available. Although we do not suggest researchers or clinicians use the correction factor as a rote multiplier, we do suggest that interested parties use them as a guide to guessing how much more aggression would have been captured. To estimate the dual-reporter level of aggression, researchers need only multiply their prevalence rate by the correction factor. We recommend that clinical researchers use the Stony Brook Clinic (i.e., Sample 1) correction factors of 1.3 (men) and 1.2 (women) for prevalence of male-to-female aggression overall and 2.4 (men) and 1.1 (women) when specifically focusing on severe acts. Note that for general occurrence, prevalence rates that are based on reports by men or women underestimate the couple-based prevalence. For severe acts, however, prevalence figures from women will capture nearly all the couple-reported acts; prevalence rates from men will tremendously underestimate the couple-based prevalence.

Furthermore, we recommend that general researchers use the correction factors from the Stony Brook Community sample and Buffalo Newlywed Study sample. (These samples produced nearly equivalent results; for ease of discussion, we refer to the correction factors produced when the samples were combined.) It is interesting that the accuracy of men's and women's reports in community samples were not strikingly different (i.e., correction factors of 1.4 for prevalence of aggression for either men or women; 1.7 and 1.4 when specifically focusing on severe acts for men and women, respectively). For example, the most widely cited prevalence rates of male-to-female aggression (Straus & Gelles, 1990) were as follows:

² Although overreporting by victims (e.g., false positives) is a theoretical possibility, no credible explanation has been offered as to why victims of partner aggression, in general, would overreport. Because of (a) the persuasive explanation that aggressors would minimize their behaviors and (b) the danger of researchers revictimizing recipients of aggression by labeling them as overreporters, we reject the tenability of this hypothesis and do not discuss it further in this article. Interview studies are needed to estimate the number of false positives attributable to measurement errors (e.g., those who misunderstood a question and reported abuse when they didn't mean to).

Table 3		
Sensitivity of the Con	flict Tactics Scale and Breakdown of Who Reports Aggression	n

	Sensitivity		Reports of aggression				
			Male	Female		Correction factor	
Study	Male report	Female report	report only (%)	report only (%)	Both report	Male report	Female report
Sample 1 (clinic)							
Any reports of violence	.78	.86	14	22	63	1.3	1.2
Severe violence	.42	.91	9	58	33	2.4	1.1
Sample 2 (community)							
Any reports of violence	.73	.69	31	27	41	1.4	1.5
Severe violence	.58	.64	36	42	22	1.7	1.6
Buffalo Newlywed Study							
Any reports of violence	.71	.76	24	29	46	1.4	1.3
Severe violence	.61	.72	28	40	33	1.7	1.4
Sample 2 + Buffalo Newlywed Study							
Any reports of violence	.72	.72	28	28	44	1.4	1.4
Severe violence	.59	.69	31	40	29	1.7	1.4

male reports — any = 10.5%, severe = 1.3%; female reports — any = 11.8%, severe = 4.9%. Using the community corrections above, we would estimate rates of 14.7% for any aggression and 2.2% for severe aggression for male respondents. This means that, had Straus and Gelles interviewed the partners of the male respondents and scored the men as aggressive if either spouse reported aggression, they may have obtained prevalences similar to the corrected rates. For women respondents, the corrected rates would be 16.5% and 7.4%.

Several issues regarding this study should be taken into consideration. First, it is possible that some spouses will report aggression by mistake (e.g., misunderstanding questions, reporting on aggression that occurred more than 1 year ago). However, given the concreteness of the questions, these false positives probably would be extremely unlikely. A future study to examine false positives could entail the following: (a) administering the CTS in its typical written form, (b) interviewing participants using an oral version of the CTS, and (c) discussing and clarifying any discrepancies between the written and oral CTS. The rate of erroneous reports on the written CTS would then be known and could be used to correct for false positives in all studies using the written CTS.

However, given the social undesirableness of

Table 4
Correction Factors From Previously Published CTS Agreement Studies

	_	Reports of aggression						
	Sensitivity		Male	Female		Correction factor		
Study	Male report	Female report	report only (%)	report only (%)	Both report	Male report	Female report	
Bohannon et al. (1995)								
Any reports of violence	.85	.57	43	15	43	1.2	1.7	
Severe violence	.65	.60	40	35	26	1.5	1.7	
Szinovacz (1983)								
Any reports of violence	.61	.65	35	39	27	1.6	1.5	
Severe violence	.36	.72	28	64	8	2.8	1.4	
Margolin (1987)	-							
Any reports of violence	.90	.74	26	10	64	1.1	1.3	
Severe violence	.63	.88	13	38	50	1.6	1.1	

Note. CTS = Conflict Tactics Scale.

spousal aggression, the number of cases in which both spouses report no violence when there really was violence (uncaught true positives measured with a hypothetical, objective measure) will be many times larger than the false positives. Because such cases are counted as true negatives, even the adjustment equations in our study provide a lower bound for estimating the actual extent of male-to-female aggression.

Second, these correction factors are useful only in estimating the overall prevalence of male-to-female aggression. Although the basic findings of this study (i.e., men and women seeking marital therapy are equally likely to mildly underreport the occurrence of aggression, whereas men substantially underreport severe aggression) may be pertinent to clinicians, one cannot apply correction factors to individual cases. The disparity between male and female correction factors for severe aggression was so great in our clinic sample (whereas it was not in the community sample) that we strongly advise all clinicians to obtain the woman's report, even if the man is the sole client.

Third, nearly all of the participants in the various studies summarized herein were married. Further work is necessary to replicate these correction factors with dating or cohabiting partners.

Fourth, all of the studies used the typical 1-year time frame for reporting aggression. Although it makes our findings applicable to other studies that use the CTS, it should be noted that (a) our couples who reported no aggression in the previous year may have some lifetime history of aggression and (b) the correction factor may be affected by the time frame and, thus, may not be generalizable to studies that use a time frame other than 1 year.

Finally, all samples (except the Buffalo Newlywed Study) included in this article were limited to either (a) volunteers responding to advertisements or (b) intact couples presenting to a marital therapy clinic. Three caveats result from the nature of the samples. First, most of the samples did not represent randomly selected groups. However, the one study of newlywed couples that was generalizable (Leonard & Senchak, 1994) produced equivalent correction factors as the volunteer Stony Brook sample couples. Thus, although we hope that our results will be replicated and extended with nationally

representative samples,3 it appears likely that our results provide generalizable correction factors for early married couples. Second, early married spouses may not follow the same reporting patterns as spouses married for longer periods (or for nonmarried partners). Thus, replication with other community samples is necessary. We provide correction factors in Table 4 from the three published CTS agreement studies that reported enough data for such calculations. The correction factors from these studies were comparable to the Stony Brook Community sample findings and the Buffalo Newlywed Study findings (except severe violence underreporting in the Szinovacz study, which was more similar to the Stony Brook clinic sample), and the variability found may have been attributable, at least in part, to the small sample sizes of the prior studies (Ns of about 100). Finally, our samples were limited to marital clinic and volunteer samples. Estimation equations for severe battering samples (i.e., from court-mandated batterers or abuse shelter samples) may be different and require future study.

References

Barling, J., O'Leary, K. D., Jouriles, E. N., Vivian, D., & MacEwen, K. E. (1987). Factor similarity of the Conflict Tactics Scale across samples, spouses, and sites: Issues and implications. *Journal of Family Violence*, 2, 37-54.

Bohannon, J. R., Dosser, D. A., & Lindley, S. E. (1995). Using couple data to determine domestic violence rates: An attempt to replicate previous work. *Violence and Victims*, 10, 133-141.

Browning, J., & Dutton, D. (1986). Assessment of wife assault with the Conflict Tactics Scale: Using couple data to quantify the differential reporting effect. *Journal of Marriage and the Family*, 48, 375–379.

Cascardi, M., Langhinrichsen, J., & Vivian, D. (1992). Marital aggression, impact, injury and health correlates for husbands and wives. Archives of Internal Medicine, 152, 1178-1184.

³ A recently published study by Szinovacz and Egley (1995) used one item (husband hit, shoved, threw things at wife) from the nationally representative National Survey of Families and Households and reported results equivalent to a correction factor of 1.6 for husbands and 1.5 for wives. Even though a 1-item measure is not equivalent to the full CTS, the findings are similar.

- Dobash, R. E., & Dobash, R. P. (1984). The nature and antecedents of violent events. *British Journal of Criminology*, 24, 269–288.
- Jouriles, E. N., & O'Leary, K. D. (1985). Interspousal reliability of reports of marital violence. *Journal of Consulting and Clinical Psychology*, 53, 419–421.
- Kurz, D. (1993). Physical assaults by husbands: A major social problem. In R. J. Gelles & D. R. Loseke (Eds.), Current controversies on family violence (pp. 88-103). Newbury Park, CA: Sage.
- Lawrence, E., Heyman, R. E., & O'Leary, K. D. (1995). Correspondence between telephone and written assessments of physical violence in marriage. *Behavior Therapy*, 26, 671-680.
- Leonard, K. E., & Senchak, M. (1994, November). Prospective study of alcohol and marriage aggression. Paper presented at the 28th Annual Meeting of the Association for the Advancement of Behavior Therapy, San Diego, CA.
- Locke, H. J., & Wallace, K. M. (1959). Short-term marital adjustment and prediction tests: Their reliability and validity. Journal of Marriage and Family Living, 21, 251-255.
- Margolin, G. (1987). The multiple forms of aggressiveness between marital partners: How do we identify them? *Journal of Marital and Family Therapy*, 13, 77–84.
- McLaughlin, I. G., Leonard, K. E., & Senchak, M. (1992). Prevalence and distribution of premarital aggression among couples applying for a marriage license. *Journal of Family Violence*, 7, 309–319.
- Navran, L. (1967). Communication and adjustment in marriage. Family Process, 6, 173–184.
- O'Brien, M. O., John, R. S., Margolin, G., & Erel, O. (1994). Reliability and diagnostic efficacy of parents' reports regarding children's exposure to marital aggression. *Violence and Victims*, 9, 45–62.
- O'Leary, K. D. (Ed). (1987). Assessment of marital discord. Hillsdale, NJ: Erlbaum.
- O'Leary, K. D., & Arias, I. (1988). Assessing agreement of reports of spouse abuse. In G. Hotaling, D. Finkelhor, J. T. Kirkpatrick, & M. A. Straus (Eds.), Family abuse and its consequences: New directions in research (pp. 218–227). Newbury Park, CA: Sage.
- O'Leary, K. D., Barling, J., Arias, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and

- stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology*, 57, 263-268.
- O'Leary, K. D., Vivian, D., & Malone, J. (1992). Assessment of physical aggression against women in marriage: The need for multimodal assessment. *Behavioral Assessment*, 14, 5-14.
- Saunders, D. (1988). Wife abuse, husband abuse, or mutual combat? In K. Yllö & M. Bograd (Eds.), Feminist perspectives on wife abuse (pp. 90-113). Newbury Park, CA: Sage.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, 38, 15–26.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence. *Journal of Marriage and the Family*, 41, 75-88.
- Straus, M. A. (1990). The Conflict Tactics Scale and its critics. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families* (pp. 49–74). New Brunswick, NJ: Transaction.
- Straus, M. A. (1993). Physical assaults by husbands: A major social problem. In R. J. Gelles & D. R. Loseke (Eds.), Current controversies on family violence (pp. 67-87). Newbury Park, CA: Sage.
- Straus, M. A., & Gelles, R. J. (1986). Societal change and change in family violence from 1975–1985 as revealed by two national studies. *Journal of Marriage and the Family*, 48, 465–479.
- Straus, M. A., & Gelles, R. J. (1990). *Physical violence in American families*. New Brunswick, NJ: Transaction.
- Straus, M. A., Gelles, R. J., & Steinmetz, S. K. (1980). Behind closed doors: Violence in the American family. New York: Anchor/Doubleday.
- Szinovacz, M. E. (1983). Using couple data as a methodological tool: The case of marital violence. Journal of Marriage and the Family, 45, 633-644.
- Szinovacz, M. E., & Egley, L. C. (1995). Comparing one-partner and couple data on sensitive marital behaviors: The case of marital violence. *Journal of Marriage and the Family*, 57, 995–1010.

Received August 9, 1996
Revision received November 22, 1996
Accepted November 24, 1996