

The Effect of Question Order on Disclosure of Intimate Partner Violence: An Experimental Test Using the Conflict Tactics Scales

Ignacio Luis Ramirez^{1,3} and Murray A. Straus²

Published online: 13 May 2006

Two plausible but contradictory approaches to question order in research on sensitive or criminal behavior are (1) that presenting the questions in a sequence corresponding to a culturally recognized behavior pattern will facilitate disclosure, and (2) that presenting questions in random order will result in more disclosure because random order disrupts response sets. The question order of the original Conflict Tactics Scales (CTS1) used the culturally recognized sequence approach, and the revised CTS (CTS2) used a modified random order. This experiment was designed to determine which of these two question orders results in more disclosure of physical and sexual assault of a dating partner. The standard version of the CTS2, which has the questions from each of the 5 scales interspersed in a slightly modified random order, was administered to every second student in a sample of 417 university students. The other half of the sample were given the same instrument but with the questions in the culturally recognized sequential order used in the CTS1. The cultural sequence order begins with the socially approved behaviors in the Negotiation scale and ends with scales measuring antisocial and the criminal behavior such as the Physical Assault scale. The results indicate that the CTS2 random order produced significantly higher disclosure rates for the scales that measure criminal behavior (Physical Assault, Injury, and Sexual Coercion) and made no difference for the other CTS2 scales (Negotiation and Psychological Aggression). Although these results suggest that the CTS2 random order is the preferred approach, reasons to treat that conclusion with caution are presented.

KEY WORDS: violence; dating partners; Conflict Tactics Scales; question order; injury; sexual coercion.

Rossi *et al.* (1983) identified the five ways in which the order of questions asked could produce different results. (1) *Saliency effect*: Question order could affect the saliency of certain topics by providing differential contexts within which responses are elicited. For example, agreement that there are “circumstances when it is ok for a husband to slap his wife’s face” is much higher if respondents are first asked the same question about a woman slapping her husband (Moore & Straus, 1995; Simon *et al.*,

2001). (2) *Redundancy effect*: Question order could affect answers if questions are not grouped together by topic. For example, questions on the same topic in different parts of an interview could make respondents wonder if they are being asked the same questions again and if so, whether there is some trick involved. (3) *Consistency effect*: A response set or pattern may occur when questions become predictable. Respondents may answer in a manner that is similar with the previous answers to portray a consistent response. (4) *Fatigue effect*: Respondents may grow tired of answering questions, especially if the questionnaire is long, resulting in an increase in unanswered questions, or a decrease in accuracy. (5) *Rapport effect*: Respondents may be nervous or hesitant in the beginning but rapport could build as the questionnaire proceeds, resulting in fewer unanswered questions and more accurate responses.

¹Department of Sociology, Texas Tech University, Lubbock, Texas 79409.

²Family Research Laboratory, University of New Hampshire, Durham, New Hampshire 03824.

³To whom correspondence should be addressed at Department of Sociology, Texas Tech University, Lubbock, Texas 79409; e-mail: l.ramirez@ttu.edu.

The question order effects identified by Rossi, Wright, and Anderson may influence the degree to which the widely used Conflict Tactics Scales (CTS) is able to secure disclosure of intimate partner violence. (In this paper the original CTS will be referred to as the CTS1, the revised CTS as the CTS2, and if both versions are the referent, CTS will be used.) The CTS1 (Straus, 1979) was designed to maximize the “rapport effect” and to minimize the “redundancy effect” by grouping questions in what will be called “cultural sequence” order. However, the cultural sequence order poses an increased risk of the “consistency effect” because it enhances the saliency of the criminal nature of the behaviors in three of the scales (Physical Assault, Injury, Sexual Coercion) and the socially desirable or noncriminal nature of the behavior in two of the scales (Negotiation and Psychological Aggression). This may make it easier to avoid disclosing the socially undesirable behaviors.

An alternative is to present the items from each of the five CTS2 scales in random order. A random order presumably disrupts response sets and may therefore result in more disclosure of the criminal behaviors. When the CTS2 (Straus *et al.*, 1996) was developed, the question order was changed to present the items in a modified random order. However, the effectiveness of this change was not empirically tested. An empirical test is also needed because the presumed benefit of a random order may be counterbalanced by a “redundancy effect” when respondents encounter questions that are similar enough to wonder if they have been asked the same question twice or more. This paper, therefore, reports the results of an experiment to compare two versions of the CTS2. One group of respondents received the CTS2 as published, with the items for each scale randomly interspersed (“CTS2 random order”). The other half of the respondents received a version that started with the most socially desirable scale, and then the remaining scales in an ascending order of severity of maltreatment of the partner (“cultural sequence order”).

LEGITIMIZING EFFECTS VERSUS RESPONSE SETS

The questions in the CTS1 were presented starting with the most socially approved conflict tactic (e.g., “Discussed an issue calmly”), and ending with the most severe of the Physical Assault scale items (e.g., “Used a knife or gun”). This sequence was based on standard principles for designing interviews, such as “the initial questions should be neither threatening nor sensitive” (Maxfield & Babbie, 1995, p. 233), and grouping questions by topic to give coherence to the interview and make it more like

a conversation. Another basis for the CTS1 sequence of questions was the findings from exploratory qualitative research done to guide the design of the CTS1. This research revealed some of the implicit cultural rules that make hitting a partner acceptable, such as having previously “tried everything,” yet the partner continues the objectionable behavior and “won’t even talk about it.” The question order of the CTS1 used this culturally recognized sequence to give participants an opportunity to show that they had “tried everything” before asking them to disclose acts of physical assault. It was intended to create a “context of legitimation” for disclosing violence against a partner (or against a child in the parent–child version of the CTS). However, despite more than 25 years of research using the CTS, including many methodological investigations, we were unable to find a study that provided empirical data on the context of legitimation theory.

On the other hand, research by Dahlstrom *et al.* (1990) has shown that randomly interspersing items asking about diverse behaviors reduce response sets and result in higher prevalence rates. Randomly interspersed items require respondents to think about each item more than would be the case if they were in groups of similar items. Respondents presumably recognize that they cannot just go down the page of a questionnaire by circling zeros (denials) because zero has the opposite meaning for questions measuring socially desirable behavior. The use of random order for items is a step to counter the “consistency effect” identified by Rossi *et al.* (1983). Random order, by providing fewer cues to identify the variables being measured, may also minimize other “demand characteristics” (Orne, 1962) that can influence how questions are answered, such as beliefs by participants about what is “really” expected of them, what the researcher is trying to prove, etc. Although this is a plausible approach, Maxfield and Babbie (1995), argue that

“...{an} attempt to overcome...{question order}... effects by randomizing the order of questions... is usually a futile effort.... A randomized set of questions will probably strike respondents as chaotic and worthless. It will be difficult to answer, moreover, since they must continually switch their attention from one topic to another. And, finally, even a randomized ordering of questions will have the {contextual} effect discussed previously, except that you will have no control over the effect.”

Maxfield and Babbie go on to point out a number of contingencies, and to recommend comparing alternative question orders, as in the experiment described in this paper.

In addition to providing data on completing principles of questionnaire design, the results of this study also have considerable practical importance because the

CTS is the most widely used method of obtaining data on intimate partner violence. Since the first use of the CTS1 in 1972, it has provided the data for approximately 400 papers, several books, and many more theses and dissertations (Yodanis *et al.*, 1997). In recent years, articles reporting findings based on use of the CTS were published at a rate of about 10 per month. An instrument that is this widely used needs to be examined carefully with a view to uncovering and correcting problems.

METHODS

Sample

Questionnaires containing the CTS2 in the published random order and in the cultural sequence order were administered during Sociology, Anthropology, and History classes at the University of Texas at El Paso and Texas Tech University during the fall 1999, spring 2000, and summer 2000 semesters. Each respondent received a booklet consisting of (1) a cover sheet explaining the purpose of the study, the participant's rights, and the name of a contact person and telephone number for those who might have questions after the test session was over. (2) Demographic questions. (3) The Personal and Relationships Profile (Straus *et al.*, 1999; Straus & Mouradian, 1999). (4) The Revised Conflict Tactics Scales (CTS2). The purpose, task demands, and rights were explained orally as well as in printed form at the beginning of each session.

Respondents were told that the questionnaire would include sensitive questions concerning attitudes, beliefs, and experiences in a relationship. They were guaranteed anonymity and confidentiality of their responses, and they were told that the session would take about an hour. All students were asked to sign a written consent form before completing their questionnaires. The actual time to finish was 30 min to 1 h. Students filled out the questionnaire at their own pace and deposited the completed (or if they chose, blank) questionnaire in a box, and left the room when they finished. A debriefing form was provided as they turned in their questionnaire. It explained the study in more detail, and provided names and telephone numbers of local mental health services and community resources, such as services for battered women.

About 600 questionnaires were distributed, and 576 of the students turned in completed questionnaires. Four hundred and seventeen met the criteria of having been in a heterosexual relationship for at least a month during the previous 12 months, and being either Mexican American or Non-Mexican White. Other ethnic groups

and homosexual relationships were not included because there were not a sufficient number to study. Because of the missing data, the number of cases varied from 417 to 347 depending on the number of variables in the multivariate analyses.

CTS2 Random Order and Cultural Sequence Order Experiment

The experimental manipulation was achieved by distributing the questionnaire with the items in the CTS2 random order to every second student in each of the classes where the questionnaire was administered. The other half of the sample responded to a version with the same items rearranged into the cultural sequence order that was used for the CTS1.

We use the term "CTS2 random order" rather than "random order," because the random order used in the CTS2 was modified by moving two of the Negotiation scale items out of random order to be the first two questions, and by moving two of the very severe violence items from the early in the set of questions to later in the instrument. These four items were moved to provide some of the context of legitimation of the CTS1 and to avoid asking difficult questions at the start. At the same time, because 90% of the 38 items remain in random order, it was believed that this would still result in disrupting the "consistency effect" that might occur with the cultural sequence question order.

Other Independent Variables

In addition to the experimentally manipulated independent variable, six other variables were included in the analysis: gender and age of the respondent, whether the respondent was the Mexican American or the Non-Mexican White, whether they were cohabiting with a partner, scores on a socioeconomic status scale and a social desirability response set scale. The questions for the first four variables and how they were coded are in Ramirez (2001).

Research that uses self-report data on crime needs to take into account the tendency of some respondents to minimize socially undesirable behavior. This study used a 13-item version of the Crowne-Marlowe social desirability response bias scale (Reynolds, 1982), as incorporated in a modified form in the Partner and Relationships Profile (Straus *et al.*, 1999). The scale measures the degree to which a respondent tends to avoid admitting the undesirable behavior, such as partner assault and other forms of crime. The items in the scale permit the respondent

to agree with or deny things that are slightly undesirable but true of everyone. For example, a person who says unequivocally that “I have never deliberately said something that hurt someone” is probably presenting themselves in a socially desirable way. The higher the social desirability score the less likely the respondent is to disclose undesirable information on the self-report survey. The correlation for this sample between the social desirability scale and reporting being assaulted by a partner was $-.17$ and $-.15$ for assaulting a partner. These are low correlations compared to the typical correlation of social desirability scales with personality tests, which is consistent with other studies of the correlation between social desirability scales and the CTS (Sugarman & Hotaling, 1996).

Socioeconomic status was measured using a 3-item scale consisting of the sum of the responses to questions on the education of the respondent’s father and mother (possible score of 1–7 for each) and family income (possible score of 1–9). The theoretical range of the resulting scale is 3–23. The alpha coefficient of reliability for the socioeconomic status scale is $.72$ (see Ramirez, 2001, for the specific questions, frequency distribution, and other statistics).

Dependent Variables

Conflict Tactics Scale Scores (CTS2)

The CTS2 has scales to measure the following five variables: physical assault, injury, sexual coercion, psychological aggression, and negotiation. See Straus *et al.* (1996) for description of the scales and data on reliability and validity. The questions in the CTS are asked in pairs. The first item in each pair asks respondents about their own behavior toward the partner and provides data on perpetration. The second item in each pair repeats the item and asks about the behavior of the partner, and therefore provides data on victimization. In addition, there are subscales for “minor” and “severe” acts. Thus, the CTS2 provides data on a total of 30-scale scores [$5 \text{ scales} \times 2 \text{ scores for types of actors (self and partner)} \times 3 \text{ scores for levels of severity (minor, severe, total)} = 30 \text{ cores}$].

Personal and Relationships Profile (PRP)

The Personal and Relationships Profile (Straus *et al.*, 1999; Straus & Mouradian, 1999) measures 22 risk factors for intimate partner violence. Some of the scales measure psychological characteristics of the respondent that are known to be associated with assaulting a partner (e.g., poor anger management skills and antisocial personality), and

some are social relationship characteristics (e.g., communication problems and dominance). There is also a social desirability response set scale to control for this source of spurious relationships. For purposes of this paper, we analyzed only scores on the Social Desirability scale, the Criminal History scale, and the Social Integration scale. They were included in the analysis only to confirm that responding to the cultural sequence or the CTS2 random order had no effect on other self-report data.

Data Analysis

Analysis of variance was used to determine if there is a significant difference in disclosure of intimate partner violence between the CTS2 random order and the cultural sequence question order versions of the CTS2. In addition to testing the main effect for question order, the model also included the gender of respondent, age of respondent (in three age groups: 18–19, 20–21, and 22 and over), Mexican American versus the Non-Mexican ethnic group, cohabiting with a partner, socioeconomic status scale quartile, and social desirability scale quartile (a $2 \times 3 \times 2 \times 2 \times 4 \times 4$ design). These six variables were included in the model as controls for possible spurious relationships and to test for interactions with question order. The ANOVAs were computed using Type III sums of squares. All effects were assessed simultaneously, with each effect adjusted for all other effects in the model. Thus, the test for each independent variable controls for the other independent variables, and the means to be presented were adjusted to control for the other independent variables. The ANOVAs were restricted to the main effects and two-way interactions because higher order interactions would have resulted in empty cells and singular variance–covariance matrices.

Effectiveness of Random Assignment of Respondents

We also used ANOVA to test the effectiveness of the random assignment of respondents to the two question order conditions. Random assignment is intended to produce groups that do not differ in respect to every possible variable. In the context of this study, this means that there should be no significant differences between the CTS2 random order and the cultural sequence order groups in respect to gender, age, percent Mexican American, percent cohabiting, or mean scores on the socioeconomic status and Social Desirability response set scales. Table I gives the statistics for these characteristics of the sample and shows that there were no significant differences.

Table I. Characteristics of Respondents in Cultural Sequence and CTS2 Random Order Groups ($N = 417$)

Characteristics	Question order		F^a
	Cultural sequence	CTS2 random	
Percent female	65	67	0.05
Age	21.7	21.9	0.93
Percent Mexican	54	61	1.34
Percent cohabiting	26	25	0.03
Socioeconomic Status Scale	12.1	12.3	0.15

^aNone were significant.

RESULTS

Exploratory Analyses

The CTS2 provides an extremely large number of scores because each of the 30 scales and subscales (see Methods section and Straus *et al.*, 1996) can be scored in four different ways: Past-Year Frequency, Past-Year Prevalence, Past-Year Chronicity, and Relationship-Length Prevalence. This makes a total of 120 different variables. To deal with this mass of data, we began by visually comparing the CTS2 random and cultural sequence groups in respect to their mean scores on all 120 variables. Two patterns came into view.

Differences Between Scales

First, we noticed that a consistent pattern of differences between the cultural sequence question order and the CTS2 random question order occurred only for the Physical Assault, Injury, and Sexual Coercion scales, whereas for the Negotiation and Psychological Aggression scales, the differences were small and not significant. A possible explanation is that physical assault, injury, and sexual coercion occur much less often than negotiation and psychological aggression. Consequently, each additional case makes a larger percentage difference. For example, suppose that in a sample of 200 cases, 100 reported acts of psychological aggression and 10 reported an attack that resulted in injury. If a certain question order results in two more instances of psychological aggression and two more cases of injury, the psychological aggression would increase by 2% and the injury by 20%.

Differences Between Scoring Methods

The second pattern occurred in respect to the scoring method. The differences were significant only for the

Past-Year Prevalence versions of the scales. A possible explanation is parallel to the explanation presented above for the differences between scales measuring relatively rare and relatively frequent behaviors. Specifically, it is possible that the scoring methods that take into account the number of times a behavior occurred increase the size of the score, and therefore make each increment resulting from a changed question order a smaller percentage change. In this connection, it is worth noting that a number of studies have found that the past-year prevalence scores provide the most consistent and theoretically meaningful results. For this reason, and also because there is little point to computing tests of significance between means that differ very little, the analysis from here on uses only the past-year prevalence scores for the Physical Assault, Injury, and Sexual Coercion scales.

Question Order Effects on CTS2 Scores

Physical Assault

The first section of Table II shows that the CTS2 random order of questions results in more disclosure of assaults on an intimate partner than when the CTS2 questions are presented in cultural sequence order. However, in the bivariate analysis, the differences are significant only for severe assaults. After adjusting for the six control variables, all the measures of physical assault show higher scores for the CTS2 random order, including both victimization (the “by partner” rows) and perpetration (the “by self” rows), and for both minor and severe assaults.

The adjusted columns show that minor assaults (e.g., slapping, throwing things in anger) were disclosed 1.7 times more often by respondents who responded to the CTS2 random order format (43% versus 25%). As expected because of the more serious nature of the crime, the rates for severe assaults (e.g., punching, kicking, attacks with objects) were much lower, but the difference between random and cultural sequence order was greater. Severe assaults were reported by about 8% of the respondents in the cultural sequence order group and by about 20% of the respondents in the random question order group, which is 2.5 times more disclosure.

Injury and Sexual Coercion

The sections of Table II for the Injury and Sexual Coercion scales show even more consistent pattern of significantly greater disclosure in response to the CTS2 random order version of the CTS2. Every one of the bivariate and the adjusted scores were significantly greater

Table II. Differences in Conflict Tactics Scale Past-Year Prevalence Scores and Personal and Relationship Profile Scores by Question Order Group

Dependent variable	Disclosure rate or mean score					
	Bivariate		<i>F</i>	Adjusted		<i>F</i>
	Cultural sequence	CTS2 random		Cultural sequence	CTS2 random	
A. Conflict Tactics Scale score disclosure rate						
Assault						
Minor by partner	29.0	32.3	0.57	25.8	43.1	6.65**
Minor by self	32.2	35.8	0.63	27.9	44.7	5.77*
Severe by partner	8.0	15.5	5.78*	7.6	20.8	7.95**
Severe by self	9.4	16.4	4.62*	8.9	20.4	5.60*
Total by partner	29.0	34.5	1.55	25.8	46.4	9.28**
Total by self	32.2	38.5	1.87	27.9	45.9	6.56*
Injury						
Minor by partner	7.6	14.8	5.51**	7.1	16.2	4.22*
Minor by self	9.1	13.9	2.44	10.4	19.7	3.97*
Severe by partner	3.3	8.1	4.53**	2.9	12.1	7.73**
Severe by self	3.3	8.6	5.30**	2.9	10.1	4.87*
Total by partner	8.1	17.5	8.49*	7.7	19.8	6.73**
Total by self	9.6	16.6	4.67**	11.0	22.1	5.23*
Sexual coercion						
Minor by partner	27.2	43.8	13.24*	26.0	42.0	5.17*
Minor by self	23.9	38.4	10.7*	22.9	36.6	4.03*
Severe by partner	6.6	12.7	4.43**	5.1	18.6	9.40**
Severe by self	5.2	9.5	2.85	4.4	13.3	5.78*
Total by partner	28.2	45.4	14.13*	26.8	46.2	7.45**
Total by self	24.4	40.5	13.06*	23.8	40.8	6.02*
B. Personal and Relationship Profile scale score						
Criminal history	2.4	2.3	0.05	2.3	2.2	0.22
Social integration	30.7	30.4	0.44	30.7	30.6	0.01
Social desirability scale	34.6	34.7	0.03	34.4	34.8	0.27

Note. The "Adjusted" column gives the means after adjustment for the age and sex of respondent, whether the respondent was Mexican American or Non-Mexican White, whether cohabiting the partner, Socioeconomic Status Scale, and Social Desirability Response Set scale.

* $p \leq .05$. ** $p \leq .01$.

when the participants responded to the CTS2 random order version of the test.

Question Order Effects on PRP Scales

The higher scores for the CTS2 random order version could be the result of some other unintended effect, such as the overall "demand characteristics" of the study (Orne, 1962). As one means of examining this issue, we compared the scores of the CTS2 random order and cultural sequence question order groups on three scales from the PRP (Straus *et al.*, 1999). The PRP is a 22-scale instrument designed for research on partner assault. Data on the reliability and validity of the PRP is given in Straus and Mouradian (1999). The PRP was the same in both versions of the questionnaire. Consequently, the question order

experimental treatment should not have affected scores on the PRP. Consistent with this expectation, Part B of Table II shows that respondents in the CTS2 random order and the cultural sequence order groups did not differ significantly in respect to the three PRP scales. These results can also be thought of as supporting the "discriminant validity" (Campbell & Fiske, 1959) of the experiment.

Interaction Effects

Only 2 of the 20 tests for interactions of question order with the demographic and control variables were significant at the .05 level. This is about the number that can be expected by chance when this many interactions are computed. Consequently, although we present these two

Table III. Significant Interactions of Question Order with Control Variables (CTS2 Past-Year Prevalence Rate)

Dependent variable	Interaction with	Disclosure rate		<i>F</i>
		Cultural sequence	CTS2 random	
Assault minor by self	Mexican American	38.5	42.6	4.00*
	Non-Mexican	17.3	46.8	
Assault total by partner	Yes cohabitation	24.8	57.0	3.99*
	No cohabitation	26.7	35.9	

* $p \leq .05$.

interactions in Table III, these results and the discussion of them must be regarded as tentative.

Interaction with Ethnic Group

The first interaction in Table III shows that the CTS2 random order elicited much more disclosure of minor assaults for the Non-Mexican Whites. Question order made much less difference for the Mexican Americans. Another way of describing the results in Table III is that the Mexican Americans disclose higher rates of minor assaults on their partners than the Non-Mexican Whites when questions are in the cultural sequence order. The disclosure rate is almost twice as much for Mexican Americans than for the Non-Mexican Whites. Perhaps, the cultural sequence established better rapport for the Mexican Americans. Another possibility is that the Mexican Americans paid more attention to each question than the Non-Mexican Whites because more than 80% used Spanish for at least some of their daily interactions. If we assume that they had less skill in English than the Non-Mexican White students, they needed to pay more attention to the questions than the Non-Mexican Whites.

To investigate this possibility, an additional ANOVA was run for the Mexican American part of the sample using a measure of acculturation to the dominant society as an independent variable. The acculturation scale is appropriate for this purpose because it is largely based on language usage (see Ramirez, 2001, for details). If differences in English language skill was the explanation for the interaction with ethnic group, the results of testing the interaction of question order with acculturation should have shown that, as acculturation increases, the differences between the cultural sequence and the CTS2 random sequence should have decreased. However, this pattern was not found, and neither the main effect nor the interaction effect for acculturation were significant.

Interaction with Cohabiting

The second interaction in Table III shows that, for students in a cohabiting relationship, the CTS2 random order results in a much higher score on the Total Assault scale, but for students who were not cohabiting, although question order makes a difference, the effect is not as strong. Perhaps the assault questions were more “sensitive” for respondents who were cohabiting, and this amplified whatever led to the random order producing higher disclosure only for the CTS scales that measure criminal behavior. However, if that were the explanation, there should also have been a significant interaction for the severe assault subscale, and that was not the case.

Importance of NonSignificant Interactions

In view of the marginal significance of the two interactions, and the unclear explanations for them, the more important result of our tests for interactions may be the nonsignificant interactions. The lack of a significant interaction with either gender or social desirability scale score are especially important. This indicates that the higher disclosure obtained using the CTS2 random order applies to both men and women, and to those who are low as well as high in tendency to present themselves in a socially desirable light.

Reliability of CTS2 Random Order and Cultural Sequence Versions

Alpha coefficients of reliability (Cronbach, 1970) were computed for the CTS2 random order and the cultural sequence question order versions of the CTS2. Only the “Total” scores for each scale were analyzed because the subscales for minor and severe behavior have so few items. Table IV shows that, in 9 out of the 10 comparisons, the alpha coefficients were larger for the cultural

Table IV. Reliability Analysis of CTS2 Past-Year Prevalence Scales by Cultural Sequence and CTS2 Question Order

	Alpha ^a	
	Cultural sequence	CTS2 random
Conflict Tactics Scale		
A. Criminal Behavior scales		
Assault by partner	0.90	0.82
Assault by respondent	0.87	0.78
Injury by partner	0.88	0.59
Injury by respondent	0.82	0.60
Sexual coercion by partner	0.78	0.65
Sexual coercion by respondent	0.79	0.62
B. Other scales		
Negotiation by partner	0.86	0.83
Negotiation by respondent	0.89	0.83
Psychological aggression by partner	0.79	0.78
Psychological aggression by respondent	0.78	0.79

^aBecause of missing data, the *N*s for the CTS2 random sequence question order group varied from 199 to 206, and the *N*s for the cultural sequence question order group varied from 198 to 217.

sequence version, although sometimes by only a small amount.

Ordinarily, a larger coefficient of reliability would be evidence in favor of the version of the scale with the larger coefficient. However, in this case, a larger coefficient may reflect the very problem that the random version of the CTS2 was designed to correct: the tendency of some respondents, when they come to the section of the scale with questions on criminal acts to avoid disclosing those behaviors. When this occurs, there is a tendency for the items in that section to have scores that are zero or very low. This increases their correlations with each other. Consequently, because the alpha coefficient of reliability is largely a function of the size of the correlations between the items in a scale, the higher correlations between items from more consistent denial of behaviors, results in a higher coefficient of reliability.

The larger alpha coefficients for the cultural sequence order are found only in part A of Table IV, which refers to scales measuring criminal behavior. This is consistent with the interpretation that the greater reliability of the cultural sequence question order reflects an internal consistency brought about by more consistently failing to disclose socially undesirable behavior.

SUMMARY AND CONCLUSION

There are plausible arguments in favor of presenting questions on sensitive, embarrassing, or criminal behav-

ior in a culturally recognized sequence versus presenting the questions in random order. Standard references on constructing questionnaires almost always recommend grouping questions into recognizable topics (e.g., Maxfield & Babbie, 1995). In addition, the qualitative research used to develop the CTS1 suggested that a question order that gives respondents an opportunity to show that they had “tried everything” and only resorted to hitting as a last resort corresponded to implicit cultural rules that make hitting a partner “understandable,” i.e., culturally recognized and tolerated. This was believed to create a “climate of legitimacy” that facilitates disclosure. The CTS1, therefore, presented the questions in “cultural sequence order,” starting with the socially acceptable conflict tactics in the Negotiation scale. By contrast, the revised Conflict Tactics Scales (CTS2) presents the items in a modified random order. A random order was believed to increase disclosure because it has the general effect of disrupting response sets, and in the case of the CTS, makes the reprehensible or criminal behaviors measured by some of the CTS scales less salient.

Both of these question-order effect theories are plausible. Hence, the need for the experiment reported in this paper to compare the effect on disclosure of intimate partner violence of presenting the questions in the quasi-random order used for the CTS2 and in the culturally recognizable sequence used in the CTS1. The experiment was conducted by administering the CTS2 with the questions in the CTS2 random order (the published version of the CTS2) to every second student in a sample of 417 students at two universities. The other half of the sample responded to the same instrument, but with the questions arranged in cultural sequence order.

The results indicate that the CTS2 random order produced significantly higher disclosure rates, but only for the scales that measure criminal behavior (Physical Assault, Injury, and Sexual Coercion) and made no difference for the other CTS2 scales (Negotiation and Psychological Aggression). Moreover, the greater disclosure of criminal and social undesirable behavior was remarkably consistent. It was found for both victimization and perpetration, for men and women, for younger and older respondents, and regardless of scores on a socioeconomic status scale, and scores on a social desirability response set scale.

These results indicate that the CTS2 random order results in a higher disclosure rate. However, the higher disclosure rate also creates a need for caution when comparing descriptive statistics such as the prevalence or frequency of physical assault with studies that used the CTS1. On the other hand, the higher disclosure rate

of the CTS2 is an indication of less measurement error. Consequently, analyses to test hypotheses about the antecedents or consequences of intimate partner violence using the CTS2 random order should yield parallel but stronger results compared to studies that tested the same hypothesis using the CTS1.

There are also reasons to treat the finding of a higher disclosure rate with the CTS2 random order with caution. First, the participants were university students reporting on a dating relationship. The applicability of these results to older persons in marital relationships is not known.

Second, the CTS2 random order produced higher disclosure only for the prevalence rate method of scoring the CTS. We suggested a possible explanation for why there were few significant differences when using the other scoring methods, but this needs to be empirically tested.

Third, the CTS is intended for personal interviews as well as self-administered questionnaires. It is possible that in a personal interview, the culturally recognized sequence effect would be more salient and could result in more complete disclosure than the CTS2 random order.

Finally, the CTS2 random order tested in this experiment departs from a pure random order by placing two questions from the Negotiation scale at the beginning and by moving two of the very severe violence items to later in the instrument. This was designed to retain some of the presumed advantage of the cultural sequence question order, while still gaining the presumed advantage of the random order. It is possible that the superior performance of the "CTS2 random order" rests on this combination rather than on random order alone. Thus, separate conclusions are needed for research design in general and for the CTS. The conclusion for the design of self-report measures is one of the qualified support for a random order approach. The conclusion for the CTS is that, despite the plausibility of the cultural sequence order, the modified random question order used for CTS2 results in more disclosure.

REFERENCES

- Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56: 81-105.
- Cronbach, L. J. (1970). *Essentials of Psychological Testing* (3rd edn.). Harper and Row, New York.
- Dahlstrom, W. G., Brooks, J. D., and Peterson, C. D. (1990). The Beck Depression Inventory: Item order and the impact of response sets. *J. Pers. Assess.* 55: 224-233.
- Maxfield, M. G., and Babbie, E. (1995). *Research Methods for Criminal Justice and Criminology*, Wadsworth, New York.
- Moore, D. W., and Straus, M. A. (1995). Tolerance for Slapping a Spouse: Higher than we thought. Paper presented at the *Annual meeting of the American Association for Public Opinion Research*, Fort Lauderdale, FL. Available at <http://PUBPAGES.UNH.EDU/~MAS2>
- Orne, M. (1962). Amount of experience in experiments as a determinant of performance in later experiments. *J. Per. Soc. Psychol.* 17: 776-783.
- Ramirez, I. L. (2001). *The Relation of Acculturation, Criminal History, and Social Integration of the Mexican Americans and the Non-Mexican Whites to Assaults on Intimate Partners*, Unpublished Dissertation, University of New Hampshire, Durham, NH.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *J. Clin. Psychol.* 38(1): 119-125.
- Rossi, P. H., Wright, J. D., and Anderson, A. B. (eds.). (1983). *Handbook of Survey Research*, Academic Press, Orlando.
- Simon, T. R., Anderson, M., Thompson, M. P., Crosby, A. E., Shelley, G., and Sacks, J. J. (2001). Attitudinal acceptance of intimate partner violence among U.S. adults. *Violence Victims* 16(2): 115-126.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The conflict tactics (CT) scales. *J. Marriage Fam.* 41: 75-88.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S., and Sugarman, D. (1999). *The Personal and Relationships Profile (PRP)*, Family Research Laboratory, University of New Hampshire, Durham, NH.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S., and Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *J. Fam. Issues* 17(3): 283-316.
- Straus, M. A., and Mouradian, V. E. (November 19, 1999). Preliminary Psychometric Data for the Personal and Relationships Profile (PRP): A multi-scale tool for clinical screening and research on partner violence. Paper presented at the *American Society of Criminology*, Toronto, Ontario.
- Sugarman, D. B., and Hotaling, G. T. (1996). Intimate violence and social desirability: A meta-analytic review. *J. Interpers. Violence* 12(2): 275-290.
- Yodanis, C., Hill, K., and Straus, M. A. (1997). *Tabular Summary of Methodological Characteristics of Research Using the Conflict Tactics Scales*, Family Research Laboratory, University of New Hampshire, Durham, NH. Available at <http://PUBPAGES.UNH.EDU/~MAS2>

Copyright of *Journal of Family Violence* is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Copyright of *Journal of Family Violence* is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.