THE STRUCTURE OF PHYSICAL VIOLENCE
IN INTIMATE RELATIONSHIPS

by

Katherine V. Regan

B.A. (Hons.), Simon Fraser University, 2000

Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of

Master of Arts

in the Department
of
Psychology

©Katherine V. Regan, 2003
SIMON FRASER UNIVERSITY
November, 2003

All rights reserved.
This work may not be reproduced in whole or part, by photocopy or
other means, without permission of the author, except for scholarly,
non-profit use, for which no further permission is required.
APPROVAL

Name: Katherine Regan
Degree: Master of Arts (Psychology)
Title of Thesis: The Structure of Physical Violence in Intimate Relationships

Examining Committee:

Chair: Dr. John McDonald
Assistant Professor

Dr. Kim Bartholomew
Senior Supervisor
Associate Professor

Dr. Cathy McFarland
Supervisor
Professor

Dr. Margaret Jackson
External Examiner
Professor
School of Criminology
Simon Fraser University

Date Approved: November 10, 2003
PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis (the title of which is shown below) to users of the Simon Fraser Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis: The Structure Of Physical Violence In Intimate Relationships

Author: Katherine Regan
MA (Psychology)

Signature: Katherine Regan

Nov. 14, 2003
ABSTRACT

Research has established that the physical violence items of the Conflict Tactics Scales (CTS) measure a single construct of violence. Little work has been done, however, to examine the relative severity of the individual items. This study therefore evaluated the structure of an expanded version of the physical violence scale of the CTS in a sample of randomly selected men and women. In total, 576 women and 547 men reported on perpetration and receipt of 14 violent acts in their relationships. Item response theory analyses indicated that the violence items were unidimensional for recipient and perpetrator reports of men’s violence, as well as for recipient and perpetrator reports of women’s violence. The items were dispersed along the underlying violence continuum, from lower to higher severity of violence, and discriminated well in this range. There was convergence between severity of acts as reported by perpetrators and recipients for men’s violence as well as for women’s violence. Results further indicated that adding together the total number of different acts endorsed to form a total (variety) score is appropriate for both men’s and women’s violence. These total scores were strongly related to likelihood of injury for receipt of violence, suggesting that the CTS can be used as a measure of severity of violence.
ACKNOWLEDGEMENTS

I would like to thank Kim Bartholomew for all her support and encouragement during this project. I greatly appreciate how she has challenged and inspired me over the years. Thanks to Cathy McFarland for her contributions to this research, in particular for compelling me to write more clearly about a fairly complex statistical technique. A thank you also to Joan Wolfe and Elizabeth Michno for always being available to answer my formatting questions.

I also wish to thank my husband, Keith, and daughter, Breannae for being there for me. Keith, I couldn't have done this without your love and unconditional support over the years. Breannae, you are a constant reminder that there is more to life than just graduate school.
# TABLE OF CONTENTS

Approval .......................................................................................................................... ii

Abstract ..................................................................................................................... iii

Acknowledgements ........................................................................................................ iv

Table of Contents .......................................................................................................... v

List of Tables ................................................................................................................ vi

List of Figures ................................................................................................................ vii

## INTRODUCTION

Violence in Intimate Relationships ........................................................................... 1
The Conflict Tactics Scales (CTS) ................................................................................ 2
Overview of the Current Study ..................................................................................... 8

## METHOD

Participants .................................................................................................................. 9
Measures ...................................................................................................................... 10
  Measure of physical violence .................................................................................... 10
  Measure of physical injury ....................................................................................... 11

## RESULTS

Men's Violence ............................................................................................................ 13
  Men's Reports of Perpetration ................................................................................ 16
  Women's Reports of Receipt .................................................................................... 16
Women's Violence ........................................................................................................ 17
  Women's Reports of Perpetration ............................................................................ 20
  Men's Reports of Receipt ......................................................................................... 20
Summary of IRT Findings ............................................................................................ 21
Composite Violence Score .......................................................................................... 23

## DISCUSSION ........................................................................................................... 25

References ................................................................................................................... 32

Appendix A: Figures Demonstrating the $a$ and $b$ Parameters .................................... 34

Appendix B: Example of IRT Analysis Applied to the CTS Physical Violence Items ......... 36
LIST OF TABLES

Table 1: Physical Violence Items Used in the Current Study Compared to the CTS2 Physical Violence Items ............................................................ 3
Table 2: Injury Items Used in the Current Study ........................................................ 12
Table 3: Estimates of IRT Parameters for Men's Violence (Men's Perpetration and Women's Receipt) Ever in the Past ....................... 14
Table 4: Estimates of IRT Parameters for Women's Violence (Women’s Perpetration and Men’s Receipt) Ever in the Past ....................... 18
Table 5: Correlations Between b Parameter Estimates for Reports of Men’s and Women’s Violence ............................................................ 23
LIST OF FIGURES

Figure 1: ICCs for Men's Violence ................................................................. 15
Figure 2: ICCs for Women's Violence............................................................ 19
INTRODUCTION

The physical assault scale of the Conflict Tactics Scales (CTS; Straus, 1979) is the most commonly used measure of partner violence. Research using the CTS typically reports rates of minor and severe violence separately rather than simply reporting rates of violence overall (e.g., Magdol et al., 1997). Accordingly, numerous studies have examined the structure of the CTS to determine whether the physical violence items form two subscales, a minor violence subscale and a severe violence subscale, as opposed to measuring a single construct of violence (e.g., Moffitt et al., 1997). Findings generally support the presence of a single underlying construct. However, these studies have typically relied on linear factor analysis and have examined simply whether the CTS violence items “hang together” well. Little work has been done to examine the severity of the individual items themselves. Consequently, the goal of the present study is to assess the ordering, in terms of severity, of the physical violence items of the CTS in a randomly selected sample of men and women. Using item response theory models, I explored whether the physical violence items of the CTS form a single continuum of severity of violence. I also examined whether the individual items represent similar or differing levels of severity for men’s and women’s violence. For example, is being slapped by one’s partner as severe, relative to other violent acts (such as pushing), for men as it is for women?

Violence in Intimate Relationships

Representative surveys using the CTS to measure physical relationship violence indicate that about 12% of couples report at least one incidence of husband-to-wife violence in the past year. About an equal proportion of couples report at least one
incidence of wife-to-husband violence in the same time period (e.g., Bohannon, Dosser & Lindley, 1995; Brinkerhoff & Lupri, 1988; Straus & Gelles, 1986). The rates of violence found by individual studies generally fluctuate around this mean. For example, Brinkerhoff and Lupri (1988), in a random sample in Calgary, found a rate of husband-to-wife violence of 10.3% and a rate of wife-to-husband violence of 13.2%. Comparably, Bohannon, Dosser and Lindley (1995), studying married military couples in the United States, found average rates of husband-to-wife violence of 13.5% and average rates of wife-to-husband violence of 12.5%.

The Conflict Tactics Scales (CTS)

The most commonly used measure of relationship violence is the CTS. The CTS frames questions about violent acts in a normative context, such that questions are presented to respondents as a general investigation of how couples resolve differences and conflict. The CTS can be used to measure perpetration or receipt of violence, or both. Although the CTS2 (the newest version of the CTS; Straus, Hamby, Boney-McCoy & Sugarman, 1996) contains a number of subscales (e.g., a psychological aggression scale), in this study, I focused exclusively on physical violence items. Although these items do not address all forms of partner abuse, such as psychological or sexual abuse, they do refer to acts that, within the criminal justice system, constitute assault.

Straus' (1979) initial analyses of the CTS indicated a single physical violence factor. When the more serious violent acts of using a knife or gun were added to the scale, however, these formed a factor separate from the less serious acts (Straus, 1979). Straus (1990a) therefore argued that the CTS physical violence items fall into ‘minor’ and ‘severe’ categories based on their presumed risk of injury (see Table 1). This division has been criticized for ignoring the intent behind an act and for assuming that
'severe' items entail a greater risk of injury than do 'minor' items (e.g., Dobash, Dobash, Wilson, & Daly, 1992; see also Straus, 1990a). For example, a man shoving his partner (a minor item) could result in greater harm than a man punching his partner (a severe item).

**Table 1:**

*Physical Violence Items Used in the Current Study Compared to the CTS2 Physical Violence Items*

<table>
<thead>
<tr>
<th>Physical violence items</th>
<th>Corresponding CTS2 items</th>
<th>CTS2 severity classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pushed or shoved</td>
<td>Pushed or shoved</td>
<td>Minor</td>
</tr>
<tr>
<td>2. Slapped</td>
<td>Slapped</td>
<td>Minor</td>
</tr>
<tr>
<td>3. Threw something that could hurt</td>
<td>Threw something that could hurt</td>
<td>Minor</td>
</tr>
<tr>
<td>4. Twisted arm or hair</td>
<td>Twisted arm or hair</td>
<td>Minor</td>
</tr>
<tr>
<td>* 5. Grabbed or held down in anger</td>
<td>Grabbed</td>
<td>Severe</td>
</tr>
<tr>
<td>* 6. Scratched or bit</td>
<td>(not in CTS2)</td>
<td></td>
</tr>
<tr>
<td>* 7. Hit with something that could hurt</td>
<td>Punched or hit with something that could hurt</td>
<td>Severe</td>
</tr>
<tr>
<td>* 8. Punched</td>
<td>Punched or hit with something that could hurt</td>
<td>Severe</td>
</tr>
<tr>
<td>9. Slammed against a wall</td>
<td>Slammed against a wall</td>
<td>Severe</td>
</tr>
<tr>
<td>10. Kicked</td>
<td>Kicked</td>
<td>Severe</td>
</tr>
<tr>
<td>11. Beat up</td>
<td>Beat up</td>
<td>Severe</td>
</tr>
<tr>
<td>12. Choked</td>
<td>Choked</td>
<td>Severe</td>
</tr>
<tr>
<td>13. Used a knife or gun</td>
<td>Used a knife or gun</td>
<td>Severe</td>
</tr>
<tr>
<td>14. Burned or scalded on purpose</td>
<td>Burned or scalded on purpose</td>
<td>Severe</td>
</tr>
</tbody>
</table>

*Note.* Items that have been changed or added in the current study are indicated by *. See Method section for details.
If any measurement instrument is to be widely used in research, how well it measures the construct of interest needs to be established. One way to understand a scale is to examine its structure. This can be done in several ways, including assessing the dimensionality of a scale using either linear factor analysis (e.g., Moffitt et al., 1997) or nonlinear factor analysis, that is item response theory (IRT) (e.g., Schafer, 1996). Although the IRT model is mathematically equivalent to the factor analytic model, it is more appropriate if dealing with dichotomous items (McDonald, 1981; Thissen, Steinberg, Pyszczynski, & Greenberg, 1983). IRT also allows for the assessment of how individual items are functioning relative to one another. For example, a math test may measure a single underlying construct of mathematical ability, but this does not mean that all items on the test are equally difficult. Some items will be easier and others will be more challenging. IRT allows for an examination of the relative difficulty of the test items.

Studies that have assessed the factor structure of the CTS in heterosexual samples using linear factor analysis have generally found that, for men's and women's violence, a single factor underlies the physical violence items, rather than minor and severe factors as proposed by Straus (1979). Unidimensional structures have been confirmed in samples of married couples (Barling, O'Leary, Jouriles, Vivian, & MacEwen, 1987), young adults (Moffitt et al., 1997), incarcerated men (Cook, 1998), and high school students in dating relationships (Cascardi, Avery-Leaf, O'Leary, & Slep, 1999). In contrast, two studies have found a two-factor structure for both men's and women's violence. In a sample of about 20,000 US military personnel, Pan, Neidig, and O'Leary (1994) found that items relating to choking, beating up, or using a knife or gun on one's partner loaded on a separate factor. Caufield and Riggs (1992), in a sample of

1 This was the case for recipient and perpetrator reports of women's violence, and for perpetrator reports of men's violence. For recipient reports of men's violence, only the item relating to using a knife or gun on one's partner loaded on a separate factor. All other physical violence items loaded on one factor.
university students reporting on their own aggression, found that an item about beating up one's partner loaded on its own factor. These studies, however, used exploratory rather than confirmatory factor analysis, and it may be that the findings would not stand up to the more stringent test of confirmatory factor analysis.

In sum, linear factor analysis of the CTS generally suggests that there is one factor underlying the physical violence items, rather than separate factors for minor and severe violence. This does not mean, however, that the physical violence items of the CTS all assess the same severity of physical violence. That is, even though all the items assess physical violence, some of the acts, such as pushing one's partner, may not be as severe as other acts, such as punching one's partner. In addition to testing the unidimensionality of a measure, IRT models show how the items in a scale are functioning relative to one another, and where on the continuum of the construct of violence, from low to high severity, each item lies (e.g., Schafer, 1996).

IRT models are mathematical expressions of the relation between an individual's responses on individual items and an underlying construct that is presumed to give rise to these responses (Crocker & Algina, 1986). An IRT analysis allows for a probabilistic rank ordering of the items comprising a measure along an underlying dimension, in this case severity of violence. An item characteristic curve (ICC) graphically represents how the probability of a response on the item varies with the level of the underlying construct, represented by the symbol \( \theta \) (theta). \( \theta \) is standardized with a mean of 0 and a standard deviation of 1. As such, scores on theta are similar to standard scores.

Two characteristics of ICCs are important, their slope and their position on the underlying construct. The slope (a), which indicates how well an item distinguishes between individual levels of the construct, is called a discrimination parameter. Larger values of a indicate steeper curves of the ICC. Items with steeper curves are useful for
distinguishing between individual levels of the latent construct (i.e., the probability of a yes response to the item, plotted on the y-axis, changes quickly from a low to a high value within a brief range of the construct). The \( a \) parameter is similar to a factor loading in linear factor analysis in that it refers to the strength of the relation between the individual items and the underlying construct. Appendix A contains a figure demonstrating the \( a \) parameter. As a further example, Figure B1 (see Appendix B) shows ICCs for the physical violence items of the CTS from a previous study (Regan, Bartholomew, Oram, & Landolt, 2002); the item numbers on this Figure refer to the items listed in Table 1. The items in Figure B1 are very discriminating, with values of \( a \) ranging from 2.40 to 9.63.

The position (\( b \), or location parameter) of the item in relation to the underlying construct indicates at what level of the underlying construct the item differentiates between a yes and a no response. The location parameter is equal to the latent score at which half the respondents answer yes to the item. When two items are located relatively close to one another, this indicates that they are assessing similar levels of the underlying construct. The \( b \) parameter provides additional information over what can be obtained from linear factor analysis: it identifies how each item is positioned relative to the other items in terms of the degree to which it assesses the underlying construct. See Appendix A for a figure demonstrating the \( b \) parameter. As a further example, the items shown in Figure B1 are located above the mean on the underlying violence continuum; values of \( b \) range between 1.12 and 2.50 (Regan et al., 2002). Ideally, a measure has a reasonable dispersion of items and the items have relatively steep slopes.

Only one study that I am aware of has applied an IRT analysis to the CTS in heterosexual relationships. Schafer (1996) used IRT to assess the ordering of the physical violence items of the CTS on severity. Using data from a sample of 533 college
students reporting on their own perpetration of violence, he found that for both men's and women's violence, the items followed a pattern from generally less severe to more severe. That is, the $b$ parameter values for the items indicated differing levels of severity. For women's violence, most items had fairly steep curves, indicating good differentiation between different levels of severity of violence. The only exception was the item referring to the use of a gun or knife, which had a fairly flat slope suggesting that it did not differentiate well between different levels of severity of violence. For women's violence, the items showed fairly good dispersion along the continuum of severity of violence. A few items appeared to be assessing a similar level of severity of violence, but the other items assessed differing levels of the underlying construct. For men's violence, the less severe items (as suggested by Straus, 1990b) had fairly steep curves and were well dispersed, but the more severe items had much flatter slopes and were clustered together at the more severe violence end of the scale. Overall, the IRT model showed fairly good fit for women's violence, indicating unidimensionality. The IRT model did not show good fit for men's violence, suggesting that the violence items may not have been assessing a single severity of violence construct. For both women's and men's violence, pushing, grabbing or shoving one's partner was the item located lowest in terms of severity.

The other study to have applied an IRT analysis to the CTS assessed violence in male same-sex relationships. Regan, Bartholomew, Oram, and Landolt (2002) used data from 284 gay and bisexual men reporting on experiences of both receipt and perpetration of violence in their same-sex relationships. For analyses of both receipt and perpetration, the CTS items were unidimensional, fairly well dispersed along the continuum of severity of violence, and had steep slopes. However, the rank ordering of the items generally did not follow the severity ordering suggested by Straus. Although a
minor violence item (*pushed or shoved*) was ranked as the least severe, some more severe violence items were ranked as less severe than other minor violence items. For example, *punched* was ranked as relatively less severe than other items, such as *slapped* and *threw something that could hurt*.

**Overview of the Current Study**

I examined the structure of a modified version of the CTS, which was used to assess physical violence in intimate relationships in a random sample of men and women living in the City of Vancouver. IRT analyses were used to determine whether the physical violence items formed a unidimensional construct in this sample for both men’s and women’s violence, as well as where and how well items discriminated along the severity of violence continuum. I examined consistencies between the relative severity of items based on perpetrator and recipient reports for both men’s and women’s violence. To confirm that this continuum represented severity of violence, I assessed the relationship between receipt of violence and degree of injury.
METHOD

The present study is based on data collected as part of the Vancouver Domestic Abuse Project (VDAP). This study was conducted in Vancouver, British Columbia, Canada, during the summer of 1997 with the purpose of learning more about domestic abuse in a community sample.

Participants

Over 1200 respondents, 614 men and 635 women, aged 19 and older from the City of Vancouver were contacted via random-digit dialling and completed a 10 to 15 minute telephone survey that assessed demographics, physical, emotional and sexual violence in intimate relationships, as well as family of origin violence. The survey was conducted in English, Mandarin, and Cantonese. The overall response rate, calculated as the number of completed interviews (1249) divided by the number of known eligible respondents (2933), was 43%.

The majority of respondents identified themselves as heterosexual (90%), 5% self-identified as gay or lesbian, 3% self-identified as bisexual, and 2% did not answer the question. Only those respondents who identified themselves as heterosexual were included in the present study. Consequently, the sample consisted of 547 men and 576 women. The mean age of the sample was 37.3 years ($SD = 14.1$, $Range = 18-87$ years). Most participants (49%) were between the ages of 19 and 34, 38% were between 35 and 54, and 13% were 55 years of age or older. The reported marital status of participants was as follows: 33% married, 10% living with a romantic partner, 43% single.

---

This project was funded by a Social Sciences and Humanities Research Council Major Grant awarded to Dr. Kim Bartholomew.
and never married, 7% divorced, 3% separated, and 4% widowed. Of the participants not married or living with a partner, 35% were in an ongoing romantic or sexual relationship at the time of the survey. The level of education for the sample was as follows: 22% of respondents had some high school, 61% had some college or university education, and 16% had some post-graduate education. The ethnic breakdown of the sample was as follows: 28% British, 23% other European, 19% Chinese/East Asian, 2% Latin, Central, or South American, and 18% Other. Ten percent of respondents did not identify a specific ethnic background. The distribution of personal income was: less than $20,000 (29%), $20,000 to $29,000 (20%), $30,000 to $39,000 (15%), $40,000 to $49,000 (10%), and over $50,000 (18%). Finally, 8% of the sample either did not know their personal income or declined to answer the question.

A comparison with the 1996 Canadian Census data revealed that, compared to the adult population of the City of Vancouver, the current sample was younger (12% more were in the 19–34-year-old category and 13% less were in the 50 or older category), had a higher education (11% fewer had some high school and 19% more had some university or college education), and had a higher personal income (27% more had an income greater than $20,000). Despite translating the survey into Cantonese and Mandarin, the Chinese ethnicity was underrepresented by approximately 10%.

Measures

Measure of physical violence

A 14-item measure was developed from the CTS (Straus, 1979) and the CTS2 (Straus et al., 1996) to assess participants' reports of receipt and perpetration of physically violent acts (see Table 1). This measure included all the items from the
original CTS, using the revised wording of the CTS2 where these items have been changed. One item, *punched or hit with something that could hurt* was expanded into two items: *punched* and *hit with something that could hurt*. Another item, *grabbed*, was changed to *grabbed or held down*, and an item relating to scratching or biting was added. For each item, respondents were first asked if they had ever engaged in a particular behaviour toward a partner. If they indicated that they had, they were asked how often they had done so in the past year. Next, they were asked if a partner had ever directed the same behaviour toward them. Again, if participants indicated that a partner had done so, they were asked how many times in the past year this had occurred. This procedure provided two reports of men’s violence, men’s reports of perpetrated violence and women’s reports of received violence, and two reports of women’s violence, women’s reports of perpetrated violence and men’s reports of received violence. The present study used only responses regarding behaviour ever in the past.³

**Measure of physical injury**

Five items assessed physical injuries that resulted from receipt of physically violent acts. Respondents who reported any receipt of violence ever in the past were asked questions about injuries sustained as a result of this violence (see Table 2). For each injury question, respondents were first asked whether this had happened to them ever in the past. If they indicated that it had, they were asked how many times it had happened in the past year. This study only used reports about injuries sustained ever in the past, to correspond with the reports regarding violent behaviour ever in the past. The

³ It would have also been possible to use responses regarding behaviour in the past year in the present study. I did not take this approach because rates of endorsement of the physical violence acts were quite low for the past-year timeframe. Low rates of endorsement make it difficult to adequately fit an IRT model to the data, particularly when coupled with a relatively large number of items and a relatively small sample size, as in the present study (Hulin, Drasgow, & Parsons, 1983).
injury items relating to injuries sustained ever in the past were combined to form an injury index ($\alpha = .74$).

Table 2:

**Injury Items Used in the Current Study**

<table>
<thead>
<tr>
<th>Injury items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Have you ever had severe injuries such as broken bones, a concussion or severe bruising?</td>
</tr>
<tr>
<td>2. Have you ever had less severe injuries such as sprains, bruises, cuts or scratches?</td>
</tr>
<tr>
<td>3. Have you ever gone to a doctor because of an injury resulting from conflict with a partner?</td>
</tr>
<tr>
<td>4. Have you ever needed to see a doctor because of an injury, but you did not go?</td>
</tr>
<tr>
<td>5. Have you ever felt physical pain that still hurt the next day because of conflict with a partner?</td>
</tr>
</tbody>
</table>
RESULTS

IRT models for the violence items were estimated separately for men's and women's violence, for both self-reports and partner-reports. Items were excluded from the analyses if they were not endorsed at all, or when only one participant had either been the recipient or perpetrator of that act. Multilog 7 (Scientific Software International, Inc., 2003) was employed to fit the two-parameter logistic model to the item responses. Along with estimates of the parameters of the model, Multilog provides a test of the goodness-of-fit of the model in the population. However, this test is not valid when the contingency table of responses is sparse, as was the case for each analysis in the present study. Consequently, I assessed the goodness-of-fit of the model, that is, the fit between the estimated ICCs and the observed data, by examining the standardized residuals, which were computed for each response option of each item. Thus, for the 14 items there were 28 standardized residuals. If the model is correct and the items assess a unidimensional underlying construct, then these statistics will have an approximately standard normal distribution with roughly 95% being less than 1.96 in magnitude (Hambleton & Swaminathan, 1985).

Men's Violence

The results of analyses for men's violence, that is, men's reports of perpetration and women's reports of receipt, are shown in Table 3 and Figure 1. The item numbers on Figure 1 correspond to the item numbers on Table 3.
<table>
<thead>
<tr>
<th>Physical violence items</th>
<th>Men's perpetration</th>
<th>Women's receipt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a (SE)</td>
<td>b (SE)</td>
</tr>
<tr>
<td>1. Pushed or shoved</td>
<td>7.04 (1.39)</td>
<td>5.31 (.72)</td>
</tr>
<tr>
<td>2. Slapped</td>
<td>2.00 (.35)</td>
<td>3.22 (.50)</td>
</tr>
<tr>
<td>3. Threw something that could hurt</td>
<td>1.97 (.62)</td>
<td>3.15 (.57)</td>
</tr>
<tr>
<td>4. Twisted arm or hair</td>
<td>2.72 (.60)</td>
<td>4.07 (.69)</td>
</tr>
<tr>
<td>5. Grabbed or held down in anger</td>
<td>2.03 (.32)</td>
<td>3.40 (.50)</td>
</tr>
<tr>
<td>6. Scratched or bit</td>
<td>2.82 (1.41)</td>
<td>3.78 (.85)</td>
</tr>
<tr>
<td>7. Hit with something that could hurt</td>
<td>1.65 (.68)</td>
<td>4.50 (1.29)</td>
</tr>
<tr>
<td>8. Punched</td>
<td>2.37 (.66)</td>
<td>5.50 (1.23)</td>
</tr>
<tr>
<td>9. Slammed against a wall</td>
<td>3.35 (.79)</td>
<td>4.18 (.76)</td>
</tr>
<tr>
<td>10. Kicked</td>
<td>2.13 (.96)</td>
<td>4.15 (1.11)</td>
</tr>
<tr>
<td>11. Beat up</td>
<td>2.93 (1.58)</td>
<td>6.96 (1.66)</td>
</tr>
<tr>
<td>12. Choked</td>
<td>2.18 (1.23)</td>
<td>3.61 (.81)</td>
</tr>
<tr>
<td>13. Used a knife or gun</td>
<td>2.79 (.77)</td>
<td></td>
</tr>
<tr>
<td>14. Burned or scalded</td>
<td></td>
<td>2.24 (.23)</td>
</tr>
</tbody>
</table>
**Figure 1:**

**ICCs for Men's Violence**

[Diagram showing the proportion endorsement across the violence continuum for men's reports of perpetration.]

[Diagram showing the proportion endorsement across the violence continuum for women's reports of receipt.]
**Men’s Reports of Perpetration**

*Used a knife or gun* and *burned or scalded on purpose* were excluded from this analysis because no men reported perpetrating these acts ever in the past. The two-parameter logistic model showed good fit for men’s perpetration of violence ever in the past, with all 24 standardized residuals less than 1.96 in magnitude.

Table 3 and Figure 1 show that the 12 items discriminated well on the underlying severity of violence continuum, with \( a \)'s ranging between 1.65 and 7.04. The items were generally ordered according to their presumed level of severity. There were, however, several exceptions to the expected ordering of items. *Grabbed or held down in anger* and *slammed against a wall*, both considered ‘severe’ items (see Table 1), had low \( b \) parameters, indicating that they discriminated best at lower levels of the continuum. In addition, *threw something that could hurt*, considered a ‘minor’ item, had a fairly high \( b \) parameter, indicating that it discriminated best at high levels of the continuum.

For this analysis, the violence items covered a fairly wide range on the underlying continuum \( (b’s = 1.00 - 3.48) \) and were fairly evenly spaced out over that range. *Pushed or shoved* was somewhat separate from the rest of the items at the lower end of the continuum, as was *hit with something that could hurt* at the upper end of the continuum.

**Women’s Reports of Receipt**

*Burned or scalded on purpose* was excluded from this analysis because no women reported being the recipient of this act ever in the past. The two-parameter logistic model showed very good fit for women’s receipt of men’s violence ever in the past, with all 26 standardized residuals less than 1.96 in magnitude.

Table 3 and Figure 1 show that the 13 items discriminated well on the underlying severity of violence continuum, with \( a \)'s ranging from 2.79 to 6.96. The items were
generally ordered according to their presumed level of severity. There were again, however, several exceptions to the expected ordering of items. *Grabbed or held down in anger* and *slammed against a wall*, considered ‘severe’ items, both had low $b$ parameters, indicating that they discriminated best at lower levels of the continuum. In addition, *threw something that could hurt*, considered a ‘minor’ item, had a fairly high $b$ parameter, indicating that it discriminated best at high levels of the continuum.

Although covering a fairly wide range on the underlying continuum ($b$’s = .88 – 2.24), the violence items near the middle and upper end of the distribution tended to cluster together. *Pushed or shoved* was fairly separate from the rest of the items at the lower end of the continuum, as was *used a knife or gun* at the upper end of the continuum. The rest of the items tended to form several smaller clusters in between the extremes, and there was one tie on the $b$ parameter: both *hit with something that could hurt* and *choked* had a value of $b = 1.87$.

**Women’s Violence**

The results of analyses for women’s violence, that is women’s perpetration and men’s receipt, are shown in Table 4 and Figure 2. The item numbers on Figure 2 correspond to the item numbers listed in Table 4.
### Table 4:

**Estimates of IRT Parameters for Women’s Violence (Women’s Perpetration and Men’s Receipt) Ever in the Past**

<table>
<thead>
<tr>
<th>Physical violence items</th>
<th>Women’s Perpetration</th>
<th>Men’s Receipt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a$ (SE)</td>
<td>$b$ (SE)</td>
</tr>
<tr>
<td>1. Pushed or shoved</td>
<td>2.67 (.36)</td>
<td>1.17 (.09)</td>
</tr>
<tr>
<td>2. Slapped</td>
<td>1.70 (.27)</td>
<td>1.66 (.17)</td>
</tr>
<tr>
<td>3. Threw something that could hurt</td>
<td>2.84 (.57)</td>
<td>1.86 (.14)</td>
</tr>
<tr>
<td>4. Twisted arm or hair</td>
<td>3.14 (.98)</td>
<td>2.54 (.29)</td>
</tr>
<tr>
<td>5. Grabbed or held down in anger</td>
<td>2.52 (.68)</td>
<td>2.21 (.24)</td>
</tr>
<tr>
<td>6. Scratched or bit</td>
<td>2.94 (.72)</td>
<td>2.04 (.17)</td>
</tr>
<tr>
<td>7. Hit with something that could hurt</td>
<td>2.28 (.77)</td>
<td>2.58 (.39)</td>
</tr>
<tr>
<td>8. Punched</td>
<td>2.80 (.49)</td>
<td>1.82 (.14)</td>
</tr>
<tr>
<td>9. Slammed against a wall</td>
<td>4.12 (1.46)</td>
<td>2.21 (.18)</td>
</tr>
<tr>
<td>10. Kicked</td>
<td>3.17 (.71)</td>
<td>2.07 (.18)</td>
</tr>
<tr>
<td>11. Beat up</td>
<td>2.55 (1.27)</td>
<td></td>
</tr>
<tr>
<td>12. Choked</td>
<td>4.52 (3.30)</td>
<td>2.73 (.43)</td>
</tr>
<tr>
<td>13. Used a knife or gun</td>
<td>2.47 (.93)</td>
<td>2.91 (.54)</td>
</tr>
<tr>
<td>14. Burned or scalded</td>
<td>2.90 (3.43)</td>
<td>3.10 (1.35)</td>
</tr>
</tbody>
</table>
Figure 2:

**ICCs for Women’s Violence**

[Graph showing the relationship between the violence continuum and proportion endorsement for women's reports of perpetration.]
Women's Reports of Perpetration

*Beat up* was excluded from this analysis because only one woman reported perpetrating this act ever in the past. The two-parameter logistic model showed good fit for women's perpetration of violence ever in the past, with all 26 standardized residuals less than 1.96 in magnitude.

Table 4 and Figure 2 show that the 13 items discriminated well on the underlying severity of violence continuum, with *a*'s ranging between 1.70 and 4.52. The items were generally ordered according to their presumed level of severity. There were again, however, several exceptions to the expected ordering of items. *Punched*, considered a 'severe' item, had a low *b* parameter, indicating that it discriminated best at lower levels of the continuum. In addition, *twisted arm or hair*, considered a 'minor' item, had a fairly high *b* parameter, indicating that it discriminated best at high levels of the continuum.

In this analysis, the violence items covered a fairly wide range on the underlying continuum (*b*'s = 1.17 – 3.10). The items near the middle of the distribution, however, tended to cluster together. *Pushed or shoved* was again fairly separate from the rest of the items at the lower end of the continuum, as was *burned or scalded on purpose* at the upper end of the continuum. The rest of the items formed several smaller clusters in between the extremes.

Men's Reports of Receipt

No items were excluded from this analysis due to lack of endorsement. The two-parameter logistic model showed good fit for men's receipt of women's violence ever in the past, with all 28 standardized residuals less than 1.96 in magnitude.

Table 4 and Figure 2 show that the 14 items discriminated well on the underlying severity of violence continuum, with *a*'s ranging from 1.41 to 4.79. The items were
generally ordered according to their presumed level of severity. There were again several exceptions to the expected ordering of items. *Punched*, considered a ‘severe’ item, had a low $b$ parameter, indicating that it discriminated best at lower levels of the continuum. In addition, *twisted arm or hair*, considered a ‘minor’ item, had a fairly high $b$ parameter, indicating that it discriminated best at high levels of the continuum.

Although covering quite a wide range on the underlying continuum ($b$'s = .74 – 4.36), the violence items clustered together somewhat. *Pushed or shoved* was again fairly separate from the rest of the items at the lower end of the continuum, as was *burned or scalded on purpose* at the upper end of the continuum. The rest of the items tended to group together in a couple of smaller clusters in between the extremes.

**Summary of IRT Findings**

In terms of discrimination, the slopes of the items in the analyses reported here were fairly steep, indicating that they discriminated well between lower and higher severity of violence. As well, the values of $b$ were generally well dispersed on the severity of violence continuum.

For all analyses, *pushed or shoved* had the lowest $b$ parameter, indicating that it was the item associated with the lowest level of severity of violence. The ordering of the items was generally as expected, although the ‘minor’ items tended to be interspersed with the ‘severe’ items in the lower and middle ranges of the continuum.

Across analyses, consistencies in terms of item ordering emerged. For both men’s and women’s violence, the items were ordered similarly according to perpetrator and recipient reports. Although the ordering of items was not identical for both reports in these analyses, the exceptions to the expected ordering of items were the same. For example, for men’s violence, *grabbed or held down in anger* and *slammed against a*
wall, both ‘severe’ items, were located lower on the severity of violence continuum than expected. Also, *threw something that could hurt*, a ‘minor’ item, was located higher on the continuum than expected.

To examine these consistencies further, I looked at similarities across form of violence (men’s and women’s) and reporter (perpetrator and recipient) by correlating the values of the $b$ parameter estimates from all four analyses. Table 5 presents the findings for the convergent and discriminant validity of men’s and women’s violence. Bolded correlations indicate convergent validity; all remaining correlations indicate discriminant validity. The convergent validity correlations indicate agreement between perpetrator and recipient reports of men’s/women’s violence about the relative severity of the violence acts. In contrast, the discriminant validity correlations assess agreement between: a) perpetrator reports of men’s and women’s violence; b) recipient reports of men’s and women’s violence; c) women’s reports of own perpetration and receipt of violence; and d) men’s reports of own perpetration and receipt of violence. I expected that the convergent validity correlations (within-men’s/women’s violence across-reporter) would be stronger than the discriminant validity correlations (across-men’s/women’s violence within-reporter and across-men’s/women’s violence across-reporter).

Consistent with predictions, the correlations between perpetrator and recipient reports for both men’s and women’s violence were very strong ($r = .94$ & $r = .80$, respectively). These correlations, indicating convergent validity, were stronger than any of the four discriminant validity correlations. That is, the strongest agreement about the relative severity of items was across different reporters within any one type of violence.
### Table 5:
**Correlations Between b Parameter Estimates for Reports of Men's and Women's Violence**

<table>
<thead>
<tr>
<th></th>
<th>Men's Violence</th>
<th>Women's Violence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1. Men's reports of perpetration</td>
<td>--</td>
<td>.94**</td>
</tr>
<tr>
<td>2. Women's reports of receipt</td>
<td>--</td>
<td>.68*</td>
</tr>
<tr>
<td>3. Women's reports of perpetration</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>4. Men's reports of receipt</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05. **p < .01.

### Composite Violence Score

Since the model fit the data well for all four analyses discussed above, indicating that the items were unidimensional, it was appropriate to form a composite of the items (Thissen et al., 1983; see also Regan et al., 2002). When dealing with a unidimensional construct, the simplest method of scoring a measure is to add up the number of items that are endorsed by any given individual to form a total score (Thissen et al., 1983). In this case, I added the number of different acts that participants reported either receiving or perpetrating, with each act receiving equal weight. Note that items that were excluded from the analyses above were also excluded when computing the total score. I then computed the model-estimated reliability (i.e., precision of measurement) of these composites using Maple 7 (Waterloo Maple Inc., 2001). The model-estimated reliability of these composites was high for both men's violence (men's reports of perpetration $r^2 = .90$; women's reports of receipt $r^2 = .96$) and women's violence (women's reports of perpetration $r^2 = .91$; men's reports of receipt $r^2 = .94$). This type of scoring (i.e., total
scores) has been used in the past with the physical violence items of the CTS, and the total scores have been referred to as variety scores (e.g., Moffitt et al., 1997).

To confirm that the variety scores do in fact represent the underlying construct of severity of violence, I looked at the relationship between each participant’s variety score and his/her Multilog-estimated score on theta, the underlying dimension of severity of violence. The relationship was linear and almost perfect for both men’s violence (men’s reports of perpetration $r = .997$; women’s reports of receipt $r = .998$) and women’s violence (women’s reports of perpetration $r = .996$; men’s reports of receipt $r = .996$).

I also looked at how well the physical violence items were functioning as a measure of severity of violence by examining the relationship between receipt of violence and degree of injury. To do this, I looked at the relationship between the injury index and individuals’ variety scores and Multilog-estimated theta scores for respondents who had been the recipients of physically violent acts. The relationship between injury and theta scores was linear and strong for both women and men ($r = .57$ & $r = .55$, respectively), as was the relationship between injury and variety scores ($r = .57$ & $r = .56$ for women and men, respectively).
DISCUSSION

This study examined the structure of the physical violence items of an expanded version of the CTS in a randomly selected sample of women and men. IRT analyses indicated that the physical violence items form a unidimensional construct for both receipt and perpetration (ever in the past) of both men’s and women’s violence. The one-factor structure does not, however, contradict the idea that the items represent different levels of severity of violence, as was demonstrated by the item-specific results. In fact, the strong associations between injury and individuals’ theta and variety scores indicate that the items form a measure of severity of violence.

In terms of individual item characteristics, some consistent findings emerged for the four analyses reported here. Across analyses, the slopes of the items were large and the distribution of the item locations was reasonable. Although more ‘severe’ items were generally located at higher levels of the continuum than more ‘minor’ items, there were several inconsistencies in the distinction between ‘minor’ and ‘severe’ items. There were also consistencies between the current and previous IRT analyses of the CTS. Notably, the placement of pushed or shoved at the low end of the continuum in all of the present analyses is consistent with expectations and findings of previous studies (Regan et al., 2002; Schafer, 1996).

Some interesting findings emerged when I looked for consistencies in item ordering across the four analyses. The ordering of items for men’s violence across different reporters (i.e., men’s perpetration and women’s receipt) followed a very similar pattern, as did the ordering for women’s violence (i.e., women’s perpetration and men’s receipt). The strong consistency in ordering is somewhat surprising, given that the violence was reported by different people (i.e., women reporting on their own receipt and
men reporting on their own perpetration and vice versa) who were not partnered and who approached the issue from different perspectives (i.e., as recipient and perpetrator). These findings suggest general agreement between women and men about severity of partner violence.

For men’s violence (i.e., men’s reports of perpetration and women’s reports of receipt), two items were located lower than expected on the severity of violence continuum: grabbed or held down in anger and slammed against a wall. It is possible that the latter item (slammed against a wall) is similar enough to pushed or shoved that it is located with the more ‘minor’ items at the lower end of the continuum. As for grabbed or held down, perhaps this is an act that occurs relatively commonly when a man is being violent, and which doesn’t generally result in serious injury. Consistent with this speculation, 48% of men who reported having been violent endorsed this act. When I examined the degree of injury reported by women who reported having been grabbed or held down, most reported no (37%) or only minor (36%) injuries. In addition, threw something that could hurt was located higher than expected on the severity of violence continuum. It is not clear to me why this item would be located higher than expected.

In contrast, for women’s violence (i.e., women’s reports of perpetration and men’s reports of receipt), punched was located lower on the severity of violence continuum than was expected. It is possible that, when perpetrated by a woman, a punch is unlikely to result in injury. To explore this further, I examined the degree of injury reported by men who had been punched. The majority reported having sustained no (49%) or only minor (40%) injuries as a result of partner violence. In addition, twisted arm or hair was located higher than expected on the severity of violence continuum. Twisting a partner’s arm or hair is an act that may only be used by women when they are engaging in a fairly serious violent incident in which they are closely involved with their
partner. When I examined the degree of injury reported by men who had had their arm or hair twisted, relatively few (13%) reported not having sustained any injury as a result of partner violence. Almost two thirds (63%) reported having sustained minor injuries, and 25% reported having sustained severe injuries. I cannot ascertain whether these injuries were in fact due to the arm or hair twisting, since all these men did report having been the recipients of more than one act of women’s violence. However, the high incidence of injury reported by men who had experienced being scratched (61% minor, 22% severe) suggests that these acts tend to take place within the context of more severe violence by women partners.

Overall, the results suggest that the severity of various violent acts may differ depending on who is being violent. In the present sample, grabbed or held down in anger and twisted arm or hair were ranked as less severe, relative to other acts measured, for men’s violence than for women’s violence. That is, being the recipient of these acts appears to be more severe, relative to other acts measured, for men than for women. As noted above, grabbed or held down is a fairly common act for men to engage in when they are being violent toward a partner, and is unlikely to result in serious injury. This may explain why this act is located lower for men’s violence than for women’s violence, relative to other acts measured. Conversely, it appears that twisted arm or hair is an act that takes place within the context of more severe women’s violence, explaining why this act is located higher, relative to other acts assessed, for women’s violence than for men’s violence.

Further, in the present sample, threw something that could hurt and punched were more severe, relative to other acts measured, for men’s violence than for women’s violence. Thus, it appears that being the recipient of these acts in a heterosexual relationship is more severe, relative to other acts measured, for women than for men.
Conversely, punching or throwing something at one’s partner does not appear to be as severe, relative to other acts, for women’s violence as for men’s violence. These differences in relative severity of items seem to make intuitive sense. As noted above, it is possible that a punch by a woman is not as likely to result in injury as is a punch by a man. Consequently, the fact that this item was located lower on the severity of violence continuum, relative to the other items, for women’s violence is perhaps not surprising. In terms of throwing something at one’s partner, it is possible that men’s greater strength, on average, makes this a more severe act when perpetrated by a man than by a woman.

This study does have some limitations. First, respondents were women and men living in a large urban center, and these findings cannot be generalized to individuals living in rural areas. Second, comparisons with Canadian Census data revealed that the community sample collected for this study was not representative of the general population. The respondents were younger, more educated, and had higher incomes than the general population. The Chinese population was also underrepresented. Therefore, generalization of these findings must be done with caution. As well, researchers in the domestic violence field have expressed concerns about relying on self-report data when examining intimate violence. The tendency of both men and women to underreport their own aggression has been well documented (see Archer, 1999). This underreporting bias is particularly relevant for researchers assessing rates of violence. It is, however, less of a concern with the type of analysis employed in the present study, because systematic underreporting of one’s own violence is unlikely to change the relative severity of the violence items.

In this study I examined the validity of an expanded version of the CTS as a measure of physical violence in heterosexual relationships by assessing the structure of the measure. I also confirmed that severity of violence as assessed by this measure is
Structure of Physical Violence

strongly related to injury level. The results of the present study suggest that the physical violence items of the CTS are appropriate to use when examining both men's and women's violence. The high reliability of all four composite measures, coupled with the strong associations between variety scores and injury, suggests that these composite measures are equally valid for men's and women's violence. The findings also indicate that there are no systematic differences in severity of acts as reported by recipients and perpetrators for both men's and women's violence. The convergence of these reports can be taken as further evidence of validity of the CTS physical violence items.

There are, however, other ways to assess validity that should be pursued in future research. For example, Straus (1990a) suggests examining inter-partner agreement about levels of violence in a relationship in order to assess concurrent validity of the CTS (for an example of this approach, see Moffitt et al., 1997). Unfortunately, it would not be possible to look at inter-partner agreement about the severity ordering of items, since IRT models can only be applied to aggregate sample data. The present study could, however, be expanded to assess validity of the CTS physical violence items in heterosexual relationships by looking at agreement about severity of items in a sample of couples reporting on their own and each others' violence. Further, the results of the present study could be replicated in clinical samples with higher levels of violence.

This study focused exclusively on physical violence. Other forms of abuse, notably psychological and sexual abuse, are important in their own right and their role in heterosexual relationships needs to be further examined. For example, research suggests that psychological and physical violence are highly correlated (e.g., Moffitt et al., 1997), and that psychological abuse may have as harmful an effect as physical violence (e.g., Follingstad, Rutledge, Berg, Hause, & Polek, 1990). It is important to keep
in mind, however, that within the criminal justice system, the physical violence acts assessed in this study would constitute assault.

The results of this study have implications for scoring the physical violence subscale of the CTS. The present study and previous research indicate that the physical violence items form a unidimensional construct. Thus, it is more appropriate to speak of physical violence overall, rather than to discuss ‘minor’ and ‘severe’ physical violence as separate constructs. The results of the IRT analyses further indicated that ‘minor’ violence items were not necessarily located lower on the severity of violence continuum than were ‘severe’ violence items. This suggests that using the traditional approach to defining ‘minor’ and ‘severe’ physical violence may misrepresent what is actually happening in violent relationships.

Given that the physical violence items of the CTS are unidimensional, it is appropriate to use a scoring method that adds together the number of different acts endorsed to form a total (variety) score. The fact that the variety scores were strongly related to individuals’ theta scores and to the likelihood of injury for receipt of violence provides support for the use of variety scores as a method of scoring the CTS.

Variety scores have two main advantages. First, equal weight is given to all abusive acts. As a result, variety scores don’t give more weight to more common acts (such as pushing), which tend to be less severe. The second advantage of variety scores is that the endorsement of more acts generally indicates greater severity since the most severe acts tend to be the least frequent. For example, in the present sample, the mean variety score for women who reported having been kicked (a ‘severe’ act) was 8 (range 2-13). Another way to demonstrate this is to examine which acts women with low and high variety scores commonly report. Of those women reporting having been the recipient of only a single act of men’s violence, the majority (40%) reported having been
pushed or shoved (the least severe act of violence). None reported having been punched or kicked, and only one reported either being slammed against a wall or choked. In contrast, of those women reporting any violence, those with higher variety scores (i.e., variety scores higher than the median of 3) reported experiencing more severe acts of violence such as being slammed against a wall (70%), punched (63%), choked (35%), and kicked (33%).

In summary, this study investigated the structure of the physical violence items of an expanded version of the CTS in heterosexual relationships. IRT analyses indicated that these items represent a single latent dimension. The analyses also indicated that the items fall on a continuum of severity of violence. The present study suggests that, although it is reasonable to use the physical violence items of the CTS to assess the severity of both men’s and women’s violence in heterosexual relationships, it may not be appropriate to discuss ‘minor’ and ‘severe’ violence as separate constructs.
REFERENCES


Appendix A:

Figures Demonstrating the $a$ and $b$ Parameters

Figure A1 illustrates how the discrimination ($a$) parameter operates. The figure depicts three ICC's, each with a different slope: $a_1 = .1$, $a_2 = 1$, and $a_3 = 100$. Looking at the figure, we can see that the first item does not discriminate effectively because the proportion of respondents answering yes is nearly the same at every level of the latent trait. However, the third item, $a_3 = 100$, discriminates very effectively between respondents with scores less than 1.45 and those with scores greater than 1.55. Finally, the second item is intermediate in terms of its discrimination.

Figure A2 illustrates how the location ($b$) parameter operates. This figure again depicts three items, but this time each has the same slope but a different location: $b_1 = .5$, $b_2 = 1.0$, and $b_3 = 1.5$. For each ICC, the vertical line indicates the latent trait score at which the proportion of respondents answering yes to that item is .5. This figure illustrates that $b$ is equal to the latent trait score at which half the respondents answer yes to the item.
Figure A1:

**ICC's demonstrating the discrimination (a) parameter**

![Graph for discrimination parameter](image)

Figure A2:

**ICC's demonstrating the location (b) parameter**

![Graph for location parameter](image)
Appendix B:

Example of IRT Analysis Applied to the

CTS Physical Violence Items

Figure B1:

ICC's for CTS Physical Violence Items

Note. Figure reprinted with permission of the author from Regan, Bartholomew, Oram, and Landolt (2002).