Measuring the Burden of Intimate Partner Violence by Sex and Sexual Identity: Results From a Random Sample in Toronto, Canada

Alexa R. Yakubovich,[1,2] [iD] **Jon Heron,[3] Nicholas Metheny,[1,4] Dionne Gesink[2] and Patricia O'Campo[1,2]**

Abstract

Debates on how sex, gender, and sexual identity relate to intimate partner violence (IPV) are longstanding. Yet the role that measurement plays in how we understand the distribution of IPV has been understudied. We investigated whether people respond differently to IPV items by sex and sexual identity and the implications this has for understanding differences in IPV burdens. Our sample was 2,412 randomly selected residents of Toronto, Canada, from the Neighborhood Effects on Health and Well-being (NEHW) study. IPV was measured using short forms of the Physical and Nonphysical Partner Abuse Scales (20 items). We evaluated the psychometric properties of this measure by sex and sexual identity. We examined whether experiences of IPV differed by sex and sexual identity (accounting for age and neighborhood clustering) and the impacts of accounting for latent structure and measurement variance. We identified differential item functioning by sex

[1]Unity Health Toronto, ON, Canada
[2]University of Toronto, ON, Canada
[3]University of Bristol, UK
[4]University of Miami, FL, USA

Corresponding Author:
Alexa R. Yakubovich, Unity Health Toronto, ON M5B 1W8, Canada.
Email: alexa.yakubovich@utoronto.ca

for six items, mostly related to nonphysical IPV (e.g., partner jealousy). Males had higher probabilities of reporting five of the six items compared to females with the same latent IPV scores. Being female and identifying as lesbian, gay, or bisexual were positively associated with experiencing IPV. However, the association between female sex and IPV was underestimated when response bias was not accounted for and outcomes were dichotomized as "any IPV." Common practices of assuming measurement invariance and dichotomizing IPV can underestimate the association between sex or gender and IPV. Researchers should continue to attend to gender-based and intersectional differences in IPV but test for measurement invariance prior to comparing groups and analyze scale (as opposed to binary) measures to account for chronicity or intensity.

Introduction

Effective prevention of intimate partner violence (IPV) requires reliable measurement of this violence (Craig et al., 2008). Yet, how to measure and interpret the burden of IPV has been the focus of decades-long debates. Measurement controversies include: the validity of different data sources (e.g., crime data vs. general surveys); measuring discrete incidents as opposed to patterns of violence; focusing on physical or sexual violence to the exclusion of psychological violence; analyzing types of IPV separately; and whether and how intent, consequence, or severity of violence should be incorporated into measurement (Heise et al., 2019; Holtzworth-Munroe & Meehan, 2004; Johnson, 1995; Walby et al., 2017; Yakubovich et al., 2019). The reporting of IPV may also be biased, so that regardless of actual experiences of IPV, internalized factors (e.g., normative expectations, socialized roles) influence people's interpretation of their experiences, their willingness to report, and ultimately how they respond to different measures (Jewkes et al., 2015). In turn, theoretical debates persist around whether and how IPV is gendered, including the extent of, and differentiation in, burdens of IPV by sexual or gender identity (Johnson, 2011; Kimmel, 2002; Peitzmeier et al., 2020; Straus, 2010). These measurement and theoretical issues are inextricably linked because different measurement and operationalizations of IPV produce different understandings of the gendered (or not) nature of this

violence. This study seeks to demonstrate this linkage and advance a way forward in these debates using a novel application of psychometric methods to a unique Canadian dataset on IPV.

Canada offers a useful case study for these debates, showing how different ways of operationalizing IPV can meaningfully impact results across the field (Dragiewicz & Dekeseredy, 2012). National statistics have historically relied on dichotomized (yes/no) items of physical and sexual IPV from the Conflict Tactics Scale (Straus et al., 1996) and dichotomized items of financial and psychological IPV (unvalidated scale) (Burczycka & Ibrahim, 2016). Analyses of these data have tended to dichotomously operationalize IPV (any IPV), finding symmetric prevalence among women and men (Lysova et al., 2019) and a higher prevalence of IPV among lesbian, gay, or bisexual-identified people compared to heterosexual-identified people (Burczycka & Ibrahim, 2016). In contrast, analyses of these same data that accounted for the severity or patterns of violence, either by using observed data on counts or impacts (Lysova et al., 2019; Romans et al., 2007) or estimating underlying or *latent* patterns of IPV (Ansara & Hindin, 2010), have found that women experienced more intense and chronic patterns of IPV compared to men. These analyses, however, have not considered sexual identity nor whether the groups being compared have systematically different responses to IPV measures. The Canadian example thus raises critical questions around measurement and analysis methods that are commonly practiced in IPV research and beyond (Potter et al., 2020; Yakubovich et al., 2019). This includes how failing to attend to gender-based and intersectional differences in the chronicity or intensity of violence, as well as willingness to report experiencing different types of violence, may distort our understanding of IPV burdens.

Psychometric analyses can inform best practices in IPV measurement and operationalization but have been underused in IPV research (Ansara & Hindin, 2010; Martin-Fernandez et al., 2019; Wareham et al., 2021; Yount et al., 2014a, 2014b). Valid between-group comparisons (e.g., by sex) in IPV burdens require that there are no systematic differences in the ways that people are responding to the measure; the same construct has to be measured in the same way among groups before we can validly compare their scores on that construct (Vandenberg & Lance, 2000). When such measurement invariance does not hold, it means that observed differences may be due to differences in the ways that people have interpreted or responded to the items (e.g., due to normative expectations or socialized roles) as opposed to real differences in experiences of violence between groups. In other words, in such cases, measurement bias confounds true group differences in latent scores.

Using psychometric analyses, we can consider overall measurement invariance based on whether items are measuring distinct dimensions of IPV and contributing to the latent construct(s) of IPV to a similar degree across groups (Vandenberg & Lance, 2000). We can also examine whether people from different groups with the same latent IPV scores respond in different ways to the items, including on an item-by-item basis—this is known as differential item functioning (DIF) and is often not considered in scale validation (Martinkova et al., 2017; Shealy & Stout, 1993). Evidence of DIF demonstrates that people are responding differently to an indicator of IPV due to some group characteristic(s) beyond real differences in the underlying construct of IPV (i.e., conditional on their latent scores). For instance, a classic example of DIF is that women are more likely to endorse having "crying spells" compared to men, independent of their underlying depression scores (e.g., due to cultural norms around the acceptability of crying among men) (Teresi et al., 2008). Not accounting for this DIF (i.e., that crying is easier for women to endorse than men with the same depression scores) in analysis will falsely underestimate men's true levels of depression compared to women.

Examining the measurement of IPV experiences by both gender and sexual identity is an important line of inquiry for future epidemiologic research on IPV given the evidence for the role of sex and gender in the severity of violence, as well as in the interpretation and manifestation of relational and personal experiences (e.g., via gender norms) (Dragiewicz & Dekeseredy, 2012; Kimmel, 2002; Romans et al., 2007; Whitehead et al., 2020; Yakubovich et al., 2019). Yet, to our knowledge, there has been no such investigation in the field using robust psychometric methods. We therefore aimed to investigate the latent structure of IPV and measurement invariance by gender (using sex as a proxy) and sexual identity, using a random sample from a diverse Canadian urban center, and draw implications for how we measure and analyze social inequities in this violence.

Methods

We used data from the Neighborhood Effects on Health and Well-being (NEHW) study in Toronto (Canada's largest city) (O'Campo et al., 2015). The study systematically randomly sampled 50 of Toronto's 140 neighborhood planning areas and then randomly selected two census tracts from each of the 50 areas. Finally, households (based on residential address) were randomly selected and screened within each census tract (mean [M] = 27 residents per census tract). One resident per household was selected and screened based on the following inclusion criteria: aged 25-64 years, able to communicate in English, and lived in the census tract for at least six months. The

final sample included 2,412 residents (response rate = 77%). Face-to-face interviews were conducted between March 2009 and June 2011. Participants provided written informed consent. The St Michael's Hospital Research Ethics Board provided ethical approval for this study.

IPV

Participants completed the Hurt/Insult/Threaten/Scream (HITS) screening tool for IPV within the last 10 years. Participants indicated whether they had a partner in the previous 10 years who had physically hurt, insulted or talked down to, threatened to harm, or screamed or cursed at them plus an added fifth item of whether a partner had restricted their actions. Those who responded affirmatively to any of the five items (37%) completed abbreviated versions of the Partner Abuse Scales (Attala et al., 1994) covering physical (e.g., my partner pushes and shoves me around violently), psychological (e.g., my partner belittles me), and sexual violence (e.g., my partner physically forces me to have sex) from a current or former partner (all 20 items summarized in the results section). Participants indicated the frequency of experiencing each item on a 6-point Likert scale (0 = never to 5 = all the time).

Covariates

Participants indicated their sex assigned at birth (0 = male or 1 = female), which we conceptualized as a proxy for gender. Risk of misclassification is low given that Toronto population estimates suggest 0.5% prevalence of transgender identity (Fleiszer et al., 2019). Participants indicated their sexual identity as 0 = heterosexual or straight, 1 = gay or lesbian, 2 = bisexual, or 3 = some other way. We operationalized sexual identity as 0 = heterosexual and 1 = lesbian, gay, bisexual, or queer (LGBQ) due to low but proportionate numbers of participants in each of the latter subgroups and common theoretical drivers of IPV across LGBQ populations (Fleiszer et al., 2019; Rolle et al., 2018). We determined age based on date of birth.

Analytic Strategy

We conducted a three-stage analytic strategy in Mplus 8.4 (Muthén & Muthén, 2017). First, to compare our data with the original scale's validation, we ran confirmatory factor analysis to determine the latent structure of the IPV measure. We compared a one-factor solution (for parsimony) to the original two-factor solution. Although some items in the Physical Partner Abuse Scale could be conceptualized as nonphysical violence (e.g., makes me afraid

for my life) and some items in the Nonphysical Scale could be conceptualized as physical/sexual violence (e.g., demands I perform sex acts I do not like), we maintained consistency with the original scales in our two-factor solution for comparison's sake. We estimated logistic models with full information maximum likelihood estimation with robust standard errors (Muthén & Muthén, 2017). We considered relative differences in model fit and the direction, magnitude, and standard error of item loadings and, for the two-factor solution, the correlation between the factors ($r > .80$ suggesting redundant factors) (Brown, 2015). Nonzero scores (i.e., "rarely" to "all the time" responses) were sparsely distributed (<5 participants per response for several items); therefore, for all latent factor analyses we used dichotomized items to prevent unreliable estimation.

Second, we evaluated measurement invariance by sex and sexual identity to establish the validity of between-group comparisons. After appraising global measurement invariance (methods and results in appendix), we considered partial measurement invariance. We fit two-parameter logistic models with robust maximum likelihood estimation. We evaluated uniform and nonuniform DIF by sex and sexual identity using a multiple indicator multiple cause model based on the Crane, van Belle, and Larson (CvBL) approach (full details in appendix) (Crane et al., 2007; Heron et al., 2012). We checked for uniform DIF (freeing item thresholds) followed by nonuniform DIF (freeing item thresholds and loadings). We then considered DIF by both sex and sexual identity by rerunning our models for sex, adding a direct effect of sexual identity on the latent IPV factor and an indirect effect of sexual identity on IPV via each item in a stepwise manner.

Finally, we analyzed the distribution of IPV by sex, sexual identity, and their interaction, adjusting for age. To determine the impact of accounting for latent structure and measurement variance, we ran three models using outcome operationalizations based on the observed data ([a] any IPV, [b] sum score of all IPV items, [c] sum score of all IPV items that did not exhibit DIF) and three models based on latent factor scores ([a] latent score using all IPV items, assuming measurement invariance, [b] latent score using all items, accounting for partial measurement invariance, [c] latent score using only non-DIF items). Models were multilevel generalized linear models accounting for clustering by census tract—distributions depended on whether the outcome was any IPV (logistic), sum scores (negative binomial), or latent scores (normal).

Results

Most participants were female and identified as heterosexual; their average age was 48 years (standard deviation [*SD*] = 11 years; Table 1). In addition, most participants were White, born in Canada, university educated, employed, making more than $75,000, married, and parents. The average household among participants was 3 people (*SD* = 2 people).

Table 1. Sample Characteristics.

	Total, *N* (%) or *M* (SD)
Sex, *N* (%)	
Female	1,390 (58%)
Male	1,022 (42%)
Sexual orientation, *N* (%)	
Heterosexual	2,278 (95%)
Gay or lesbian	79 (3%)
Bisexual	32 (1%)
Other[a]	8 (<1%)
Age (years), *M* (SD)	48.4 (10.6)
Race/ethnicity, *N* (%)	
White	1,355 (57%)
Arabic/West Asian	21 (1%)
Asian	269 (11%)
Black/Caribbean	151 (6%)
Indigenous	25 (1%)
Jewish	153 (6%)
Latin American	61 (3%)
Other	356 (15%)
Immigrant status, *N* (%)	
Born in Canada	1,496 (62%)
>10 years in Canada	769 (32%)
<10 years in Canada	143 (6%)
Education, *N* (%)	
High school or less	490 (20%)
Diploma or bachelor's	1,390 (58%)
Graduate or professional degree	531 (22%)

Table 1. Continued

	Total, *N* (%) or *M* (*SD*)
Employment status, *N* (%)	
Not working	731 (30%
Currently working or student	1,679 (70%)
Income, *N* (%)	
<$75,000	674 (28%)
>$75,000	1,728 (72%)
Marital status, *N* (%)	
Never married	500 (20%)
Separated, divorced, or widowed	370 (16%)
Cohabiting	220 (9%)
Married	1,322 (55%)
Any children, *N* (%)	
No	857 (36%)
Yes	1,555 (64%)
Household size (people), *M* (*SD*)	3.1 (1.5)

Note. [a]Free text specified as bicurious (*n* = 1), heteroflexible (*n* = 1), open (*n* = 1), queer (*n* = 3), or did not want to specify (*n* = 1).

In total, 32% of participants reported experiencing any IPV (*M* = 1.9 acts of IPV, *SD* = 3.7; Table 2). Items from the nonphysical scale (with the exception of being called ugly, coerced to not work or go to school, and coerced to not socialize with family) had higher prevalence estimates across groups compared to items from the physical scale. There were descriptive differences in IPV by sex and sexual identity. A considerably larger proportion of female LGBQ-identified participants (51%) reported experiencing IPV compared to all other groups (31%-36%). Male heterosexual-identified participants reported experiencing fewer acts of IPV (*M* = 1.6) compared to all other groups (*M* = 2.1-2.7). Finally, most physical scale items were more common among females than males, apart from dangerous objects being thrown (which was least common among female LGBQ-identified participants) and being physically forced to have sex or badly hurt during sex (which was least common among male heterosexual-identified participants).

Table 2. Prevalence of Each IPV Item by Sex and Sexual Identity.

	Total, N = 2,349	Heterosexual or Straight, N = 2,216		LGBQ[a], N = 119	
		Females, N = 1,292	Males, N = 924	Females, N = 51	Males, N = 64
Items from Partner Abuse Scale: Nonphysical (Attala et al., 1994)					
My partner belittles me	606 (26%)	351 (27%)	215 (23%)	19 (37%)	19 (29%)
My partner does not want me to have any friends	253 (11%)	134 (10%)	101 (11%)	8 (16%)	10 (16%)
My partner tells me I am ugly and unattractive	143 (6%)	92 (7%)	43 (5%)	4 (8%)	4 (6%)
My partner insults or shames me in front of others	317 (14%)	187 (14%)	108 (12%)	10 (20%)	12 (19%)
My partner is stingy in giving me money	242 (10%)	170 (13%)	59 (6%)	6 (12%)	7 (11%)
My partner belittles me intellectually	303 (13%)	197 (15%)	84 (9%)	10 (20%)	12 (19%)
My partner feels that I should not work or go to school	110 (5%)	76 (6%)	30 (3%)	3 (6%)	1 (2%)
My partner does not want me to socialize with my family	230 (10%)	129 (10%)	93 (10%)	4 (8%)	4 (6%)
My partner screams and yells at me	597 (25%)	324 (25%)	235 (26%)	19 (37%)	15 (23%)
My partner has no respect for my feelings	515 (22%)	316 (24%)	168 (18%)	14 (27%)	17 (27%)
My partner acts like a bully toward me	368 (16%)	230 (18%)	113 (12%)	14 (27%)	11 (17%)

(continued)

Table 2. Continued

	Total, N = 2,349	Heterosexual or Straight, N = 2,216		LGBQ[a], N = 119	
		Females, N = 1,292	Males, N = 924	Females, N = 51	Males, N = 64
My partner is often jealous	438 (19%)	234 (18%)	171 (19%)	16 (31%)	16 (25%)
My partner demands that I perform sex acts that I do not enjoy or like	164 (7%)	114 (9%)	33 (4%)	8 (16%)	9 (14%)
Items from Partner Abuse Scale: Physical (Attala et al., 1994)					
My partner pushes and shoves me around violently	105 (4%)	71 (6%)	29 (3%)	4 (8%)	1 (2%)
My partner makes me afraid for my life	80 (3%)	67 (5%)	8 (1%)	5 (10%)	0 (0%)
My partner throws dangerous objects at me	89 (4%)	48 (4%)	37 (4%)	1 (2%)	3 (5%)
My partner tries to suffocate me with pillows, towels, or other objects	14 (1%)	13 (1%)	0 (0%)	3 (6%)	0 (0%)
My partner has broken one or more of my bones	12 (1%)	8 (1%)	2 (<1%)	1 (2%)	0 (0%)
My partner physically forces me to have sex	61 (3%)	44 (3%)	10 (1%)	4 (8%)	3 (5%)
My partner badly hurts me while we are having sex	38 (2%)	29 (2%)	6 (1%)	1 (2%)	1 (2%)
Overall IPV					
Any IPV	760 (32%)	426 (33%)	283 (31%)	26 (51%)	23 (36%)
Sum IPV score, M (SD) [range: 0-20]	1.9 (3.7)	2.1 (4.0)	1.6 (3.2)	2.7 (3.9)	2.3 (3.6)

Note. [a]Gay or lesbian, bisexual, or some other way (free text specified as bicurious [1], heteroflexible [1], open [1], queer [3], or did not want to specify [1]).

Evaluating Latent Factor Structure and Measurement Invariance

A single factor solution (assuming that all items load onto a single latent construct) fit the data very well, with items loading highly onto the latent factor (standardized loadings: 0.89-0.96) with small standard errors (0.01-0.03). The two-factor solution (assuming separate latent constructs for physical and nonphysical IPV) only marginally improved model fit relative to the one-factor model (e.g., the Bayesian Information Criterion, BIC, decreased) and the correlation between the two factors was very high ($r = .95$). We therefore adopted the one-factor solution in all further analyses for the sake of parsimony. One item had zero variation for males (no males reported that their partner had tried to suffocate them) and was therefore dropped from further analyses. The appendix provides additional analyses of the latent structure and global measurement invariance between groups.

After establishing the unidimensionality of the scale, we explored allowing partial measurement invariance through an analysis of DIF by sex, sexual identity, and their interaction. Table 3 shows the item discriminations (loadings) and difficulties (related to thresholds) by sex in the final two-parameter logistic model. Across males and females, all items were highly discriminating and most items were more difficult (i.e., required a higher latent IPV score) to report that they occurred; this indicates that the scale is most accurate in its measurement of more severely unhealthy relationships. There were six items that showed uniform DIF by sex (differences in items difficulties; Figure 1). Males, compared to females with the same values on the underlying IPV construct, had a higher probability of reporting five items: partner does not want me to have friends, does not want me to socialize with my family, screams or yells at me, is often jealous, and throws dangerous objects at me (Panels A-E, Figure 1). Females had a higher probability of reporting having had a partner who has made them afraid for their life compared to males (Panel F). No items showed nonuniform DIF (differences in discriminations and difficulties) by sex. We did not find evidence for uniform or nonuniform DIF by sexual identity nor differences in the sex-DIF results between heterosexual and LGBQ-identified participants.

IPV on Sex and Sexual Identity: The Impacts of Accounting for Latent Structure and DIF

After establishing the latent structure and partial measurement invariance by sex and sexual identity, we analyzed the association between IPV and each of sex, sexual identity, and their interaction (adjusting for age and neighborhood clustering). Across all models (Table 4), we found no evidence for an
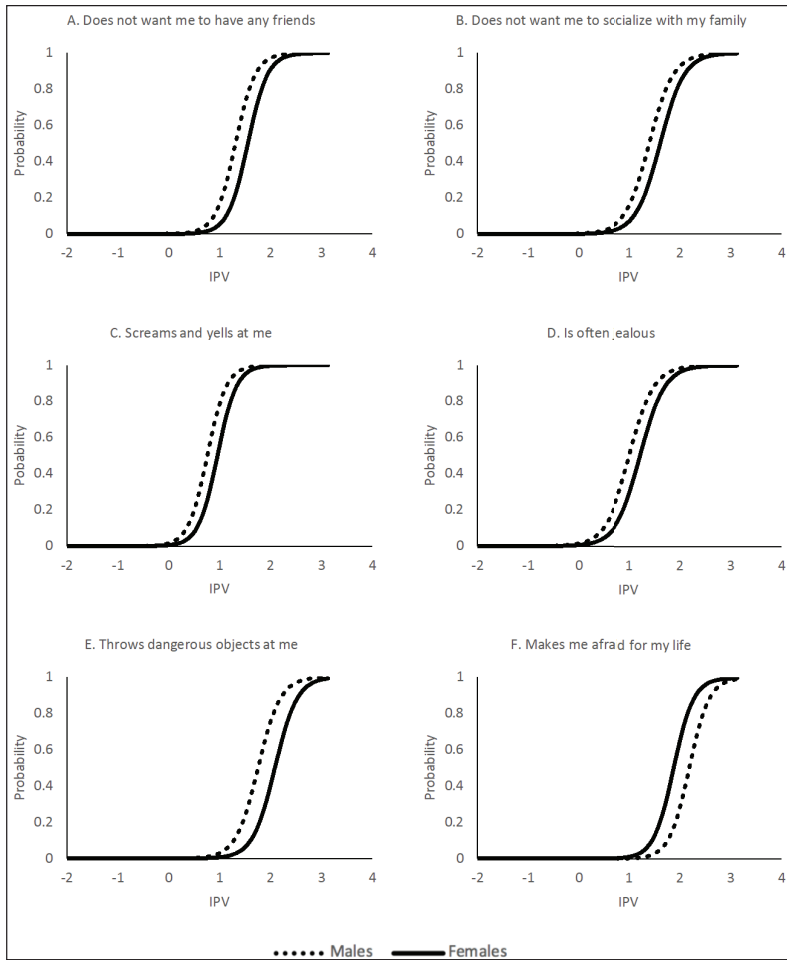
Figure 1. DIF by sex.

Note. Item characteristic curves for the six items that showed differential item functioning (DIF) by sex, estimated from the final two parameter logistic model. Panels A-E indicate uniform DIF favoring males (dotted line): males found these items easier to endorse than females across all levels of the latent IPV construct (males' item characteristic curves shifted to the left of females'). Panel F indicates uniform DIF favoring females (solid line): females found this item easier to endorse than males across all levels of the latent IPV construct (females' item characteristic curve shifted to the left of males').

Table 3. Item Discriminations and Difficulties From the Two-Parameter Logistic IRT Model by Sex.

	Females		Males	
	Discrimination	Difficulty	Discrimination	Difficulty
My partner belittles me	6.26	0.86	6.26	0.86
My partner does not want me to have any friends[a]	5.18	1.55	5.18	1.31
My partner tells me I am ugly and unattractive	4.90	1.73	4.90	1.73
My partner insults or shames me in front of others	4.54	1.34	4.54	1.34
My partner is stingy in giving me money	3.96	1.51	3.96	1.51
My partner belittles me intellectually	4.65	1.37	4.65	1.37
My partner feels that I should not work or go to school	3.02	2.03	3.02	2.03
My partner does not want me to socialize with my family[a]	4.25	1.62	4.25	1.41
My partner screams and yells at me[a]	5.41	0.96	5.41	0.76
My partner has no respect for my feelings	7.54	0.99	7.54	0.99
My partner acts like a bully toward me	6.91	1.22	6.91	1.22
My partner is often jealous[a]	4.15	1.23	4.15	1.01
My partner demands that I perform sex acts that I do not enjoy or like	3.38	1.78	3.38	1.78
My partner pushes and shoves me around violently	4.72	1.87	4.72	1.87

(continued)

Table 3. Continued

	Females		Males	
	Discrimination	Difficulty	Discrimination	Difficulty
My partner makes me afraid for my life[a]	5.17	1.89	5.17	2.19
My partner throws dangerous objects at me[a]	4.60	2.09	4.60	1.76
My partner has broken one or more of my bones	3.37	2.90	3.37	2.90
My partner physically forces me to have sex	4.28	2.13	4.28	2.13
My partner badly hurts me while we are having sex	3.97	2.34	3.97	2.34

Note. Item discriminations correspond to item loadings in traditional factor analysis and indicate how strongly correlated the item is to the latent construct of intimate partner violence. High item discriminations indicate that the item tells us more information about participants' total IPV scores (but only across the range of IPV scores for which the item is most informative). Item difficulties correspond to item thresholds and indicate the latent IPV score at which participants are more likely to endorse the item. Higher item difficulties indicate that participants need to have a higher latent IPV score before they will endorse the item.
[a]Item showed differential item functioning (DIF) by sex.

interaction between sex and sexual identity. We did, however, observe differences in the main effects of sex and sexual identity depending on how we defined the outcome. First, sexual identity but not sex was associated with the odds of experiencing any IPV: LGBQ-identified participants had higher odds of experiencing any IPV than heterosexual-identified participants (Table 4, model A). In contrast, for all other models, sex was positively associated with IPV scores: females had higher IPV scores than males (models B-F). Second, accounting for DIF led to larger estimates of the association between sex and IPV, regardless of whether this was by excluding DIF items (model C vs. model B; model F vs. model D) or by adjusting for DIF (model E vs. model D). Third, the associations between IPV scores and each of sex and sexual identity were smaller and more precise when the underlying latent structure of the IPV scores was accounted for (models D-F vs. models B-C).

Table 4. The Impact of Accounting for Latent Structure and DIF on the Estimated Associations Between IPV and Each of Sex, Sexual Identity, and Their Interaction.

	Observed Scores			Estimated Latent Scores		
	A. Any IPV[a]	B. Sum Score of IPV (All Items)[b]	C. Sum Score of IPV (Non-DIF Items)[b]	D. Latent Score of IPV (All Items)[c]	E. Latent Score of IPV (All Items, DIF Adjustment)[c]	F. Latent Score IPV (Non-DIF Ite...
Main effect models						
Female sex (vs. male sex)	0.13 (−0.05, 0.32)	0.24 (0.03, 0.46)	0.37 (0.15, 0.58)	0.14 (0.03, 0.25)	0.21 (0.10, 0.32)	0.19 (0.08, 0.3[
LGBQ (vs. heterosexual)	0.49 (0.08, 0.89)	0.23 (−0.26, 0.72)	0.31 (−0.17, 0.79)	0.21 (0.03, 0.40)	0.21 (0.02, 0.40)	0.22 (0.03, 0.4[
Interaction effect models						
Sex*sexual identity	0.18 (−0.64, 0.99)	−0.26 (−1.25, 0.73)	−0.44 (−1.40, 0.53)	0.04 (−0.33, 0.41)	0.04 (−0.33, 0.41)	−0.12 (−0.50, 0.[

Note. IPV is intimate partner violence. DIF is differential item functioning. LGBQ is lesbian, gay, bisexual, or queer. All models were multilevel (accounti for clustering by census tract) and included age as a covariate (which was consistently negatively associated with IPV). Main effect models refer to mod that only include the main effects of sex and sexual identity. Interaction effect models refer to models that include sex, sexual identity, and their interactio Results shown are the nonexponentiated point estimate and the 95% confidence interval. All models assume measurement invariance unless otherwi noted. As shown in Table 3 and Figure 1, six items showed evidence of DIF; therefore, DIF adjustment refers to analyses that accounted for DIF in these items, whereas non-DIF items refer to analyses that used an outcome variable created only from the 13 items without evidence of DIF.
[a]Multilevel logistic regression.
[b]Multilevel negative binomial regression. Negative binomial regression was used as opposed to Poisson regression as the overdispersion factor was statistically significant.
[c]Multilevel linear regression using maximum likelihood estimation with robust standard errors.

Discussion

This study found that being female and identifying as LGBQ were each positively associated with experiencing IPV in a random sample in Canada's largest city. However, the size and precision of these associations depended on whether measurement variance was accounted for. To our knowledge, only one previous study has examined measurement invariance of IPV by sex/gender in Canada, which identified response bias but did not detail the findings (Ansara & Hindin, 2010). Internationally, studies have demonstrated measurement variance by sex/gender in attitudes toward IPV (Yount et al., 2014a, 2014b) and, more recently, IPV perpetration (Wareham et al., 2021). Our findings extend this evidence base to experiences of IPV.

Six of 20 IPV items showed response bias; holding their IPV scores constant, participants' responses to these items systematically differed as a function of their sex. Independent of their latent IPV scores, males had a higher probability of reporting four items related to nonphysical IPV (e.g., partner jealousy) as well as partners throwing dangerous objects. These results map onto prior qualitative studies showing that women more often than men view nonphysical acts of IPV as controlling behaviors central to dynamics of violence (O'Campo et al., 2016). To the extent that male participants interpreted these items less severely than females, this may explain their greater likelihood of reporting these items. An additional study that compared DIF in women's responses regarding their own use of violence and that of their male partners' found that women were more likely to report that they threw objects as opposed to experiencing this (Reichenheim et al., 2007). These reporting differences may reflect gendered notions around the acceptability or severity of different forms of violence. In contrast, conditioning on latent IPV, female participants in our study were more likely to report that their partners made them fear for their lives compared to males. Previous studies have shown fear of partner to be one of the strongest gender differences in IPV self-reports (Yakubovich et al., 2019). Our results indicate that this is beyond gender differences in underlying IPV levels alone; rather, there are reporting differences on this item that may be exacerbated by, for instance, gender-based impacts of IPV (women are more likely to experience negative consequences of IPV) or constructs of masculinity (e.g., as strong and aggressive) versus femininity (e.g., as weak and subservient) (Yakubovich et al., 2019). Our study showed that ignoring item reporting biases by sex underestimated females' higher IPV scores compared to males—demonstrating the importance of testing and accounting for measurement variance prior to drawing group-based comparisons in IPV, and other epidemiologic research.

We did not identify DIF by sexual identity, however, there were a small number of LGBQ-identified participants (5%). Although a higher proportion than Canadian population estimates (Burczycka & Ibrahim, 2016), this likely limited our statistical power to identify item response bias. Low power may have also impacted identifying DIF in rarer IPV items (e.g., suffocation). Replication should be attempted in larger samples as well as in other contexts and with other IPV measures. Although alternative IPV measures have been explored with sexual and gender minority populations (Stephenson & Finneran, 2013), analyses of social inequities in IPV burdens benefit from IPV measures that exhibit measurement invariance across groups (to compare "like with like") (Martin-Fernandez et al., 2019; Shealy & Stout, 1993; Yount et al., 2014a). It would be valuable for future qualitative research to explicitly explore reasons for DIF by gender/sex and other social identities (Martinkova et al., 2017).

Despite low statistical power, our study extends previous work demonstrating the importance of sexual identity to understanding distributions of IPV (Badenes-Ribera et al., 2016; Finneran & Stephenson, 2013; Kimmes et al., 2017; Rolle et al., 2018). Prior research has predominantly been based in the United States, with a reliance on convenience or purposive samples. Only a small number of quantitative studies have demonstrated a similar or higher prevalence of IPV among same-sex couples or LGBQ-identified people compared to heterosexual populations in Canada (Burczycka & Ibrahim, 2016; Whitehead et al., 2020). Using a more robust measurement and analytic approach, our study adds to these to highlight the importance of nuancing conceptions of IPV to consider intersections between gender, sexual, and other social identities, including experiences of misogyny, colonialism, heterosexism, poverty, racism, transphobia, and other forms of structural violence (Peitzmeier et al., 2020; Ristock et al., 2017; Stark & Hester, 2019). This should be coupled with the continued broadening of services (as needed) to the unique needs of people experiencing IPV across all social locations and intersections, while recognizing that experiences and consequences of IPV are not equally distributed (Furman et al., 2017; Gingras, 2018).

Our results further demonstrate that dichotomizing IPV scales masks valuable information regarding the severity of this violence, even when these data are skewed toward zero (which they usually are) (Martin-Fernandez et al., 2019; Yakubovich et al., 2019). There is a theoretically meaningful difference between experiencing "any IPV" (which sexual identity appears more strongly predictive of) and "overlapping or more frequent acts of IPV" (which sex/gender appears more strongly predictive of): this matters to how we understand clinical burdens and design intervention strategies (Ford-Gilboe

et al., 2016; Heise et al., 2019). Analysis of the measure's latent structure indicated a single IPV construct that varies in chronicity and intensity, of which severe physical (including sexual) violence congregates at the upper end of the distribution, with certain forms of nonphysical violence congregating more toward the lower end. This further explains our analyses of dichotomized "any IPV": sexual identity was more predictive than sex because LGBQ-identified participants (especially females) tended to experience more nonphysical violence; whereas sex was more predictive of count/continuous measures because females tended to experience more physical violence than males. These results also demonstrate the information loss that can result from analyzing types of IPV separately without considering underlying structural relationships—perhaps especially when working with short-form scales (Ford-Gilboe et al., 2016; Potter et al., 2020; Yakubovich et al., 2019).

In addition to the need to test for generalizability, discussed above, study limitations include only having access to participants' sex assigned at birth (gender identity was not measured in the NEHW study) (O'Campo et al., 2015). Although a viable proxy for gender given our use of a random community sample and low prevalence estimates of transgender identity in the source population (Fleiszer et al., 2019), future research should examine measurement invariance across the diversity of gender identities—particularly in light of a growing body of research demonstrating the disproportionate burden of IPV among transgender populations (Peitzmeier et al., 2020). We also do not know the sex or gender of participants' partners or the consequences of this violence, which would be useful to consider in future research.

Summary of Implications

This study offers important context to debates around sex and gender symmetry in IPV. Females and LGBQ-identifying participants experienced more IPV, but the association between sex and IPV was underestimated when measurement variance in this construct was not accounted for. The operationalization of IPV had further consequences for understanding the distribution of IPV, with analyses of summative or latent scores clarifying the relationships between sex, sexual identity, and different severities and types of IPV in contrast to analyses of "any IPV" alone. We therefore recommend that, prior to drawing group-based comparisons in IPV, and other complex epidemiologic outcomes, researchers establish at least partial measurement invariance between groups and avoid creating outcome typologies or dichotomies without investigating latent structure. Moreover, the extent to which existing studies on sex- and gender-based differences in IPV have followed these recommended measurement and analytic methods should be interrogated to

gauge the potential role of bias in their conclusions on sex or gender symmetry. This is particularly important in contexts, like Canada, where national statistics on IPV have often used gender neutral frameworks that combine and dichotomize experiences of IPV. Practice and policy should aim to prevent and respond to IPV among all groups of people; however, strategies must remain responsive to the unequal distributions in this violence by sex, gender, sexual identity, and other social factors.

Declaration of Conflicting Interests

Funding

ORCID iD

Alexa R. Yakubovich https://orcid.org/0000-0003-4216-4305

References

Ansara, D. L., & Hindin, M. J. (2010). Exploring gender differences in the patterns of intimate partner violence in Canada: A latent class approach. *Journal of Epidemiology & Community Health*, *64*(10), 849-854. https://doi.org/10.1136/jech.2009.095208

Attala, J. M., Hudson, W. W., & McSweeney, M. (1994). A partial validation of two short-form partner abuse scales. *Women & Health*, *21*(2/3), 125-139.

Badenes-Ribera, L., Bonilla-Campos, A., Frias-Navarro, D., Pons-Salvador, G., & Monterde, I. B. H. (2016). Intimate partner violence in self-identified lesbians: A systematic review of its prevalence and correlates. *Trauma Violence Abuse*, *17*(3), 284-297. https://doi.org/10.1177/1524838015584363

Brown, T. A. (2015). Confirmatory factor analysis for applied research. In D. A. Kenny (Ed.), *Methodology in the social sciences*. The Guilford Press.

Burczycka, M., & Ibrahim, D. (2016). *Family violence in Canada: A statistical profile, 2014*. Canadian Centre for Justice Statistics. https://www150.statcan.gc.ca/n1/pub/85-002-x/2016001/article/14303-eng.htm

Craig, P., Dieppe, P., Macintyre, S., Mitchie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: The new Medical Research Council guidance. *British Medical Journal*, *337*, a1655.

Crane, P. K., Gibbons, L. E., Ocepek-Welikson, K., Cook, K., Cella, D., Narasimhalu, K., Hays, R. D., & Teresi, J. A. (2007). A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research*, *16*(Suppl. 1), 69-84. https://doi.org/10.1007/s11136-007-9185-5

Dragiewicz, M., & Dekeseredy, W. S. (2012). Claims about women's use of non-fatal force in intimate relationships: A contextual review of Canadian research. *Violence Against Women*, *18*(9), 1008-1026. https://doi.org/10.1177/1077801212460754

Finneran, C., & Stephenson, R. (2013). Intimate partner violence among men who have sex with men: A systematic review. *Trauma Violence Abuse*, *14*(2), 168-185. https://doi.org/10.1177/1524838012470034

Fleiszer, P., Ahmed, S., Ansara, D. L., Arthur, A., Blot, S., Collier, S., Corson, L., Gournis, E., Heng, J., Near, K., Rilkoff, H., & Tsirlin, D. (2019). *T.O. health check: An overview of Toronto's population health status*. Toronto Public Health. https://www.toronto.ca/wp-content/uploads/2019/11/92ef-TOHealth-Check_2019.pdf

Ford-Gilboe, M., Wathen, C. N., Varcoe, C., MacMillan, H. L., Scott-Storey, K., Mantler, T., Hegaty, K., & Perrin, N. (2016). Development of a brief measure of intimate partner violence experiences: The Composite Abuse Scale (Revised)-Short Form (CASR-SF). *BMJ Open*, *6*(12), e012824. https://doi.org/10.1136/bmjopen-2016-012824

Furman, E., Barata, P., Wilson, C., & Fante-Coleman, T. (2017). It's a gap in awareness: Exploring service provision for LGBTQ2S survivors of intimate partner violence in Ontario, Canada. *Journal of Gay & Lesbian Social Services*, *29*(4), 362-377. https://doi.org/10.1080/10538720.2017.1365672

Gingras, C. I. (2018). *The impact of heteronormativity: Barriers in seeking domestic violence services for LGBTQ+ survivors* [MSc]. Southern Connecticut State University.

Heise, L., Pallitto, C., García-Moreno, C., & Clark, C. J. (2019). Measuring psychological abuse by intimate partners: Constructing a cross-cultural indicator for the sustainable development goals. *SSM—Population Health*, *9*, 100377. https://doi.org/10.1016/j.ssmph.2019.100377

Heron, J., Crane, C., Gunnell, D., Lewis, G., Evans, J., & Williams, J. M. (2012). 40,000 memories in young teenagers: Psychometric properties of the autobiographical memory test in a UK cohort study. *Memory*, *20*(3), 300-320. https://doi.org/10.1080/09658211.2012.656846

Holtzworth-Munroe, A., & Meehan, J. C. (2004). Typologies of men who are maritally violent: Scientific and clinical implications. *Journal of Interpersonal Violence*, *19*(12), 1369-1389. https://doi.org/10.1177/0886260504269693

Jewkes, R., Flood, M., & Lang, J. (2015). From work with men and boys to changes of social norms and reduction of inequities in gender relations: A conceptual shift in prevention of violence against women and girls. *The Lancet*, *385*(9977), 1580-1589. https://doi.org/10.1016/s0140-6736(14)61683-4

Johnson, M. P. (1995). Patriarchal terrorism and common couple violence: Two forms of violence against women. *Journal of Marriage and Family*, *57*(2), 283-294.

Johnson, M. P. (2011). Gender and types of intimate partner violence: A response to an anti-feminist literature review. *Aggression and Violent Behavior*, *16*(4), 289-296. https://doi.org/10.1016/j.avb.2011.04.006

Kimmel, M. S. (2002). "Gender symmetry" in domestic violence: A substantive and methodological research review. *Violence Against Women*, *8*(11), 1332-1363. https://doi.org/10.1177/107780102237407

Kimmes, J. G., Mallory, A. B., Spencer, C., Beck, A. R., Cafferky, B., & Stith, S. M. (2017). A meta-analysis of risk markers for intimate partner violence in same-sex relationships. *Trauma, Violence, & Abuse*, *20*(3), 374-384. https://doi.org/10.1177/1524838017708784

Lysova, A., Dim, E., & Dutton, D. G. (2019). Prevalence and consequences of intimate partner violence in Canada as meaured by the national victimization survey. *Partner Abuse*, *10*(2). http://doi.org/10.1891/1946-6560.10.2.199

Martin-Fernandez, M., Gracia, E., & Lila, M. (2019). Psychological intimate partner violence against women in the European Union: A cross-national invariance study. *BMC Public Health*, *19*(1), 1739. https://doi.org/10.1186/s12889-019-7998-0

Martinkova, P., Drabinova, A., Liaw, Y. L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking equity: Why differential item functioning analysis should be a routine part of developing conceptual assessments. *CBE Life Sciences Education*, *16*(2). https://doi.org/10.1187/cbe.16-10-0307

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

O'Campo, P., Wheaton, B., Nisenbaum, R., Glazier, R. H., Dunn, J. R., & Chambers, C. (2015). The neighbourhood effects on health and well-being (NEHW) study. *Health & Place*, *31*, 65-74. https://doi.org/10.1016/j.healthplace.2014.11.001

O'Campo, P., Zhang, Y. J., Omand, M., Velonis, A., Yonas, M., Minh, A., Cyriac, A., Ahmad, F., & Smylie, J. (2016). Conceptualization of intimate partner violence: Exploring gender differences using concept mapping. *Journal of Family Violence*, *32*(3), 367-382. https://doi.org/10.1007/s10896-016-9830-2

Peitzmeier, S. M., Malik, M., Kattari, S. K., Marrow, E., Stephenson, R., Agenor, M., & Reisner, S. L. (2020). Intimate partner violence in transgender populations: Systematic review and meta-analysis of prevalence and correlates. *American Journal of Public Health*, *110*(9), e1-e14. https://doi.org/10.2105/AJPH.2020.305774

Potter, L. C., Morris, R. G., Hegarty, K., Garcia-Moreno, C., & Feder, G. (2020). Categories and health impacts of intimate partner violence in the World Health Organization multi-country study on women's health and domestic violence. *International Journal of Epidemiology*, *50*(2), 652-662. https://doi.org/10.1093/ije/dyaa220

Reichenheim, M. E., Klein, R., & Moraes, C. L. (2007). Assessing the physical violence component of the Revised Conflict Tactics Scales when used in heterosexual couples: An item response theory analysis. *Cadernos de Saude Publica*, *23*(1), 53-62.

Ristock, J., Zoccole, A., Passante, L., & Potskin, J. (2017). Impacts of colonization on Indigenous two-spirit/LGBTQ Canadians' experiences of migra-

tion, mobility and relationship violence. *Sexualities*, *22*(5-6), 67-84. https://doi.org/10.1177/1363460716681474

Rolle, L., Giardina, G., Caldarera, A. M., Gerino, E., & Brustia, P. (2018). When intimate partner violence meets same sex couples: A review of same sex intimate partner violence. *Front Psychol*, *9*, 1506. https://doi.org/10.3389/fpsyg.2018.01506

Romans, S., Forte, T., Cohen, M. M., Du Mont, J., & Hyman, I. (2007). Who is most at risk for intimate partner violence? A Canadian population-based study. *Journal of Interpersonal Violence*, *22*(12), 1495-1514.

Shealy, R. T., & Stout, W. F. (1993). An item response theory model for test bias and differential test functioning. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning*. Lawrence Erlbaum Associates, Inc.

Stark, E., & Hester, M. (2019). Coercive control: Update and review. *Violence Against Women*, *25*(1), 81-104.

Stephenson, R., & Finneran, C. (2013). The IPV-GBM scale: A new scale to measure intimate partner violence among gay and bisexual men. *PLoS One*, *8*(6), e62592. https://doi.org/10.1371/journal.pone.0062592

Straus, M. A. (2010). Thirty years of denying the evidence on gender symmetry in partner violence: Implications for prevention and treatment. *Partner Abuse*, *1*(3), 332-362.

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2) development and preliminary psychometric data. *Journal of Family Issues*, *17*(3), 283-316.

Teresi, J. A., Ramirez, M., Lai, J.-S., & Silver, S. (2008). Occurrences and sources of differential item functioning (DIF) in patient-reported outcome measures: Description of DIF methods, and review of measures of depression, quality of life and general health. *Psychology Science*, *50*(4), 538.

Vandenberg, R. J., & Lance, C. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4-70.

Walby, S., Towers, J., Balderston, S., Corradi, C., Francis, B., Heiskanen, M., Helweg-Larsen, K., Margaert, L., Olive, P., Palmer, E., Stockl, H., & Strid, S. (2017). *The concept and measurement of violence against women and men*. Policy Press.

Wareham, J., Wagers, S. M., Rodriguez, L. M., & Neighbors, C. (2021). An exploration of measurement invariance across sex in intimate partner violence perpetration. *Victims & Offenders*. https://doi.org/10.1080/15564886.2021.1934211

Whitehead, J., Dawson, M. R., & Hotton, T. (2020). Same-sex intimate partner violence in Canada: Prevalence, characteristics, and types of incidents reported to police services. *Journal of Interpersonal Violence*. https://doi.org/10.1177%2F0886260519897342

Yakubovich, A. R., Heron, J., Feder, G., Fraser, A., & Humphreys, D. K. (2019). Intimate partner violence victimisation in early adulthood: Psychometric properties of a new measure and gender differences in the Avon longitudinal study of parents and children. *BMJ Open*, *9*(3), e025621. https://doi.org/10.1136/bmjopen-2018-025621

Yount, K. M., VanderEnde, K., Zureick-Brown, S., Anh, H. T., Schuler, S. R., & Minh, T. H. (2014a). Measuring attitudes about intimate partner violence against women: The ATT-IPV Scale. *Demography*, *51*(4), 1551-1572. https://doi.org/10.1007/s13524-014-0297-6

Yount, K. M., VanderEnde, K., Zureick-Brown, S., Minh, T. H., Schuler, S. R., & Anh, H. T. (2014b). Measuring attitudes about women's recourse after exposure to intimate partner violence: The ATT-RECOURSE scale. *Journal of Interpersonal Violence*, *29*(9), 1579-1605. https://doi.org/10.1177/0886260513511536

Author Biographies

Alexa R. Yakubovich, PhD, is a CIHR Health System Impact postdoctoral fellow and CIHR-CMHC postdoctoral fellow at the Centre for Urban Health Solutions at Unity Health Toronto and the University of Toronto. She received her PhD in evidence-based social intervention and policy evaluation from the University of Oxford. Alexa's research focuses on the social and structural determinants of interpersonal violence and related health inequities and the design and evaluation of preventive interventions.

Jon Heron, PhD, is a senior research fellow in the School of Social and Community Medicine at the University of Bristol. He is a statistician with over 15 years of experience working with the Avon Longitudinal Study of Parents and Children. His methodological interests include longitudinal modeling and latent variable analysis.

Nicholas Metheny, PhD, is a registered nurse, assistant professor of nursing and health studies at the University of Miami in Coral Gables, Florida, and affiliate scientist at the MAP Centre for Urban Health Solutions at St Michael's Hospital in Toronto. He is an emerging scholar focused on prevention strategies for IPV in structurally disadvantaged populations, including women and 2SLGBTQ+ persons, and the evaluation of social policy and interventions.

Dionne Gesink, PhD, is a professor and associate dean of academic affairs in the Dalla Lana School of Public Health at the University of Toronto. She has decades of experience leading community and indigenous-partnered research with expertise in sexual and mental health, sex and gender minority population well-being, mixed-methods, and integrated knowledge translation.

Patricia O'Campo, PhD, is a Tier 1 Canada Research Chair in Population Health Intervention Research, the Executive Director of the Li Ka Shing Knowledge Institute at St Michael's Hospital, and Professor at the Dalla Lana School of Public Health. She is an internationally recognized leader in research on urban social inequalities and the structural determinants of health, including population health interventions, intersectoral solutions, and policy research.