Patterns of Women's exposure to psychological violence: A global examination of low- and middle-income countries

Cari Jo Clark [a,*], Irina Bergenfeld [a], Yuk Fai Cheong [b], Hector Najera [c], LynnMarie Sardinha [d], Claudia García-Moreno [d], Lori Heise [e,f]

[a] *Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA*
[b] *Department of Psychology, Emory University, Atlanta, GA, USA*
[c] *Programme for Development Studies. National Autonomous University of Mexico PUED-UNAM, Mexico*
[d] *The UNDP-UNFPA-UNICEF-WHO-World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), Department of Sexual and Reproductive Health and Research, World Health Organization, Geneva, Switzerland*
[e] *Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA*
[f] *Prevention Collaborative[1], USA*

A B S T R A C T

Introduction: Under Sustainable Development Goal 5, prevalence of intimate partner violence (IPV) is a globally reportable indicator. There is a lack of consensus on how to measure and report psychological IPV, affecting prevalence estimates and cross-country comparability. We examine similarities and differences in the patterning of women's experiences of psychological abuse in low- and middle-income countries (LMICs) to inform common cut points.
Methods: Data include 13,452 ever-partnered women from six LMICs participating in the WHO multi-country study on women's health and domestic violence against women and 306,101 from 47 LMICs participating in the Demographic and Health Surveys. A confirmatory latent class analysis (LCA) approach was applied to identify the optimal class structure using the 3 DHS and 4 WHO psychological IPV items, assessed the impact of physical and sexual IPV on class structure, and tested class generalizability across countries. We validated the three-class solution by regressing the classes on physical IPV, sexual IPV, controlling behaviors, and injury due to domestic violence. We used item response theory (IRT) methods to assess item-level characteristics of the items.
Results: Analysis confirmed the three-class structure in most countries. Addition of physical and sexual IPV did not change overall class structure or improve discrimination or homogeneity of the items. The three-class structure was invariant within most WHO-classified regions. Operationalized classes informed by the LCA resulted in prevalences of roughly 90% low-to-no class, 7% moderate-intensity class, and 3% high-intensity class. Classes showed convergent validity with all outcomes tested. IRT analysis revealed good discriminations but substantial information overlaps over a narrow range of the latent psychological violence construct.
Conclusions: This study confirms the three-class pattern but suggests some differences across countries. and regions. We suggest cut points distinguishing violent from non-violent acts and demarcating levels of severity for future study. Findings offer evidence-based guidance to rectify challenges.

1. Introduction

For the first time, the prevalence of intimate partner violence (IPV) in the past 12 months is a globally agreed reportable indicator in Goal 5 (Gender Equality and Women and Girls' Empowerment) of the Sustainable Development Goals (SDGs) (United Nations, 2015b). This is due

to decades of advocacy, and more recently, to the growing evidence on its extent; its health, economic, and social consequences; and its preventability. Notably, indicator 5.2.1 acknowledges all three major forms of IPV (physical, sexual, and psychological). This indicator was categorized by the Inter-agency and Expert Group on SDG Indicators as conceptually clear, having internationally established methodology and

standards (e.g. (United Nations Statistical Office, 2014)), and being regularly produced by at least 50% of countries where the indicator is relevant (United Nations, 2023). However, the conceptual clarity of psychological IPV significantly lags that of physical and to some extent sexual IPV. The authors of the original WHO Multi-country Study on Women's Health and Domestic Violence (WHO), one of the earliest and largest multi-country investigations of IPV, refrained from publishing prevalence estimates of psychological violence due to concerns about a lack of cross-cultural consensus on what constitutes this phenomenon (Claudia García-Moreno et al., 2005). Two global reviews of IPV prevalence data have subsequently been produced, neither of which provided estimates for psychological violence citing the same concerns (Devries et al., 2013; Sardinha et al., 2022). To date, no cross-country assessment of the scope and meaning of psychological violence has been conducted and investigators remain unclear as to how to define a case of psychological violence in a globally comparable way, often applying different cut-points based on frequency, number or types of acts, or combining measures of psychological violence with measures of male controlling behaviors or economic violence. Further, psychological IPV often co-occurs with physical and sexual IPV, making it difficult to ascertain what is being measured when separate estimates of psychological, physical, and sexual IPV are presented alone. As psychological violence is among the most frequently experienced forms of IPV by women in many countries (Halim et al., 2018; Martín-Fernández et al., 2019; Semahegn et al., 2019), understanding this phenomenon is essential to the accurate and consistent reporting across countries of progress toward SDG 5.

In 2019, Heise and colleagues published a re-analysis of data from the original WHO Multi-country Study—a probability-based household survey conducted between 2000 and 2004 among women at 15 sites in 10 countries (Bangladesh, Brazil, Ethiopia, Japan, Namibia, Peru, Samoa, Serbia and Montenegro, Thailand, United Republic of Tanzania) to ascertain if there were underlying patterns of psychologically abusive experiences that differentiated extreme patterns of exposure, which may be more universally recognized as IPV, from those that are a consequence of poor conflict resolution and communication (Heise et al., 2019a). In this prior analysis, a three-class model of psychological abuse was identified based on frequency (did not occur, occurred once or twice, or many times) of four items measuring insulting, belittlement or humiliation, intimidation or scaring, and threats of harm to the woman or others she cared about. This model was characterized into classes defined by high-intensity (N = 7%), moderate-intensity (N = 10%), and little to no exposure to psychological abuse (83%). The high-intensity class was defined as having a high probability of endorsing at least one item frequently. The moderate-intensity class consisted of some exposure across items, with insults being the only item appearing many times. Little to no exposure consisted of almost zero exposure to any items. This model showed a graded relationship with controlling behaviors and almost every measured heath outcome except for pain (Heise et al., 2019a).

While the study advanced our understanding of the global patterning of women's experiences of violence, it was restricted to 10 countries with data nearly two decades old. Additionally, the study did not investigate the potential impact of simultaneously considering physical and sexual IPV on the number and nature of the classes. In the present study, we use a much larger pool of more contemporary data on psychologically abusive acts to: 1) test the patterning of women's experiences of psychological abuse in six regions of the world based on WHO classifications (Africa, Americas, Eastern Mediterranean, Europe, Southeast Asia, Western Pacific) representing 325,627 women (306101 from the DHS data and 19526 from the original WHO dataset); 2) assess the generalizability of the latent classes to identify the extent to which the number of classes, the probability of an item being endorsed conditional on class, and the probability of class membership are consistent across regions; 3) test the potential impact of including physical and sexual IPV on the structure of the classes that emerge; 4) test the

construct validity of the resultant class structure or structures with measures of controlling behaviors, physical and sexual IPV; assess the items' ability to measure the latent construct of psychological IPV within each country in the sample.

2. Material and methods

2.1. Data source, inclusion and exclusion criteria

Demographic and Health Surveys (DHS) are nationally representative household surveys that provide data for a wide range of health and population indicators. In each country, a proportion of households are randomly selected for the Domestic Violence module, and only one ever-partnered woman aged 15–49 per selected household is administered this module. Women missing data on all three psychological IPV items were excluded. For this analysis, we included the most recently completed DHS from all countries with a Domestic Violence module in their individual-level DHS questionnaire. We retained only countries with all three standard DHS psychological violence items for a total sample of 47 countries across the 6 WHO regions. For analyses using the full set of DHS IPV items (psychological, physical, and sexual), three countries lacking a full set of physical or sexual IPV items were excluded. Country, years of survey, DHS module version, and number of participants are documented in Supplemental Table 1a. Across the 47 countries, 306,101 women with IPV data were included in this analysis. Women not currently in a domestic partnership were included, as they may still have experienced abuse within the past year from a former partner.

For the analysis, we also used the data from the WHO, Multi-country study (Garcia-Moreno et al., 2006) a cross-sectional population-based household survey on domestic violence and health. In total, the survey was fielded in 15 sites in 10 countries, most using a two-stage cluster sampling design to choose households. One female occupant between the ages of 15–49 per household was asked to participate, except in Japan where the age range was 18–49 years. Household response rates were between 59.8% and 99.6%, while individual response rates were between 85.0% and 99.7% (World Health Organization, 2005) Overall, 24097 women completed the survey. The sample was restricted to ever-partnered women in countries in which data was collected from an urban and a rural area, or the entire country, leaving 13452 for the analysis (Supplemental Table 1b). Countries with only urban or only rural samples were excluded, as these would not be representative of their respective countries.

2.1.1. Variable measurement

Variables used in the LCA and IRT models are presented in Supplemental Table 2. DHS items were limited to those measures of IPV collected in the DHS questionnaire version 5 or later, which measured **psychological violence** using three items, physical IPV using seven items, and sexual IPV using two items. Participants were asked whether they had ever experienced each behavior from a husband or partner, and those that responded "yes" were asked whether the behavior occurred never, sometimes, or often within the past 12 months. Psychological violence in the past year was thus modeled trichotomously. WHO data used four items to measure psychological violence, three of which were conceptually similar (if not worded identically) to the DHS data with one additional item measuring scaring/intimidation. Participants were asked whether they had ever experienced each behavior from a current or former husband or partner, and those that responded "yes" were asked whether the behavior occurred never, once or twice, a few times, or many times within the past 12 months. The "once or twice" and "a few times" categories were combined into a single category for parsimony, to parallel DHS response options, and to address sparse data across the cells, in addition to the prior empirical finding that high frequency violent experiences are associated with worse health outcomes (Clark et al., 2019; Heise et al., 2019a).

To test whether exposure to physical and sexual IPV impacted the overall patterning of psychological abuse, we included exposure to these forms of IPV in LCA. We combined sexual IPV items into a single measure and modeled it dichotomously (ever/never). Based on prior research suggesting that violence patterning may differ based on severity (Clark et al., 2019; Spencer et al., 2016), physical IPV items were combined into a single measure, modeled trichotomously based on item severity (any severe/moderate only/none). Items considered severe were designated as such in the original DHS data, based on the Revised Conflict Tactics Scales (Straus et al., 1996). Psychological IPV is associated with and predictive of other types of IPV (Salis et al., 2014); therefore, we hypothesized that risk of other IPV-related outcomes would be associated with class membership. To test the utility and validity of our classification system, we used the dichotomous sexual IPV and trichotomous physical IPV variables described above. We also created a trichotomous measure of controlling behaviors based on the number of acts reported (none/1 or 2/3 or more) and a dichotomous measure of injury due to IPV (ever/never).

2.2. Analysis

Descriptive analyses were performed in STATA 16 (StataCorp, 2019), and latent class and item response theory models were estimated in Mplus 8.0 (Muthén & Muthén, 2018). Cluster variables and individual sampling weights from the original datasets were used in all analyses, and final weights for pooled analyses were constructed using individual weights from the original data and the sampling rates of women 15–49 years old in each country based on these populations in year the data was originally collected (https://www.who.int/data/maternal-newborn -child-adolescent/indicator-explorer-new/mca/women-of-reproductive -age-(15-49-years))-population-(thousands)).

2.2.1. Latent class analysis

We sought to examine and compare patterning of women's experience of psychological IPV across LMICs; therefore, latent class analysis (LCA), a latent, person-centered analytic approach whereby individuals can be classified into subgroups (or classes) based on their responses to measured variables, was deemed most appropriate (Collins & Lanza, 2010). Following a standard approach to similarity testing with LCA (Collins & Lanza, 2010; Finch, 2015; Schmiege et al., 2018), we ran LCA with one through four classes using maximum likelihood estimation. The range of classes explored reflects the range of potential classes based on prior research and the number plausible given the number of items. The best fitting model for the individual country and pooled regional samples was assessed by examining differences in the chi square difference test, consistent Akaike Information Criterion (CAIC), and the sample-adjusted Bayesian Information Criterion (SABIC) with attention to the principles of parsimony and interpretability (Collins & Lanza, 2010). Entropy and posterior probabilities were examined to ascertain the model's ability to classify the study population into patterns. If the model fit indices diverged (Finch, 2015), the CAIC was taken as the most reliable of the information criteria (Zhang et al., 2018). We also considered the interpretability of the class structure and its consistency across countries. The resultant classes were then examined for homogeneity (i.e., how similar the women are within classes, designated by conditional probabilities >0.7 or <0.3, and for discrimination (how statistically different the classes are from one another designated by odds ratios of class membership >5.0 or <0.2) (Ulbricht et al., 2018).

2.2.2. Latent class generalizability testing

Once patterns of psychological IPV were identified in each country, we aimed to assess how similar these patterns were across countries within the same region. For the regions that demonstrated a similar number of latent classes (configural similarity), we subsequently tested whether the item response category probabilities, conditional on class, were similarly invariant across countries (structural similarity). We compared two models, one in which the item response probabilities were allowed to vary and one in which they were constrained to be equal across countries. We also tested the similarity of the relative size of classes across countries by first fitting an unconstrained model, followed by a model in which the prevalences are constrained (distributional similarity). Model fit testing relied on the CAIC, SABIC, and the likelihood ratio test (Perra, 2020) to test whether the imposition of the constraints led to a significant worsening of model fit.

2.2.3. Operationalization of classes and validity testing

To establish the validity of the predominant class structure and its utility for identifying individuals at risk for other IPV-related outcomes, we created a set of decision rules for operationalizing class membership based on the class prevalences and conditional probabilities observed in the LCA. We based these operationalized class rules on both theoretical considerations, including item severity, and on the observed proportions of individuals endorsing each item in a given class. Using multinomial logistic regression, we then estimated relative risk ratios of four theoretically related outcomes (physical IPV, sexual IPV, controlling behaviors, and injury due to IPV) for the operationalized moderate and high intensity classes relative to the low class. We accounted for the complex sampling design at the cluster and stratum level in all models, and incorporated sampling weights where available.

2.2.4. Item response theory modeling

We performed item response theory (IRT) analysis to assess the items' ability to measure the latent construct of psychological IPV within each country in the sample. We used graded response models, suitable for ordinal response categories (Toland, 2014) to describe item-level characteristics in the WHO and DHS datasets. We estimated each item's discrimination, which describes how well the item differentiates between individuals at different levels of psychological IPV, and two difficulty parameters, which describe the level of psychological IPV at which 50% of individuals will endorse a "never" response versus a "sometimes" response, and a "sometimes" response versus an "often" response, respectively (Bichi & Talib, 2018). We also visualized item information curves and total information curves for each item set. Taken together, these provided insight into which items were contributing the most information on the psychological IPV construct across countries and the range of the latent construct being captured by each of the items and the full item set.

2.2.5. Ethical considerations

The WHO study, which inspired this analysis, adhered to the best practices and ethical and safety standards for collecting data on violence against women (Claudia García-Moreno et al., 2005; Health W. D. o. G. W. a., 2001). Additionally, WHO's ethical review group, the WHO Secretariat Committee for Research in Human Subjects, the local partner institution, and as needed the national ethical review boards from each site gave permission for the study to occur (Claudia García-Moreno et al., 2005). The DHS data used in this analysis were also collected using best practices. The standard procedures and questionnaires for all DHS surveys were reviewed and approved by institutional and host country institutional review boards. (USAID) The ethical standards for the DHS surveys are based on informed and voluntary participation, privacy and confidentiality during data collection and data processing and biomarker referral, treatment, and counseling. (USAID) This secondary analysis was submitted for review by the Emory University IRB and was determined to not be human subjects' research.

3. Results

3.1. Item prevalence

The proportion of women experiencing any listed act of psychological violence within the last year ranged from 9.31% in Western Pacific

region to 22.94% in Eastern Mediterranean region in the DHS dataset (Supplemental Table 1). In all cases, the most reported act of psychological abuse was "insulted you or made you feel bad about yourself" (13.14%) and the least frequent act was "threatened to hurt or harm you or someone you care about" (5.70%) (Supplemental Table 2). In the WHO dataset (Supplemental Table 3) the prevalence of women experiencing at least one of the psychologically abusive behaviors within the 12 months preceding the interview was between 12.29% (Samoa) and 28.60% (Tanzania). Like the DHS data, being insulted was the most reported act of violence within the last 12 months (19.98%) whereas threats of harm within the past 12 months was similarly the least frequently reported act (4.48%) (Supplemental Table 4).

3.2. Individual country LCA

Results of individual country LCA are presented in Table 1. In analyses with the three psychological violence items, all 47 DHS countries showed evidence of a three-class model, although this evidence was more ambiguous for two countries in the Africa region (Chad and Comoros), two countries in the Europe region (Kyrgyz Republic and Tajikistan), two countries in the Eastern Mediterranean region (Pakistan and Afghanistan), and both countries in the southeast Asia region (Philippines and Cambodia), for which there was also some evidence supporting a two-class model as best fitting the data. In the 44-country analyses that included physical and sexual IPV, 82% (n = 36) of countries best fit a three-class model, 7% (n = 3) a two-class model, and 11% (n = 5) a four-class model. Evidence of three-class solutions was ambiguous for eight countries in the Africa region (Burundi, Gabon, Gambia, Kenya, Nigeria, Sierra Leone, Togo, and Zambia), one country in the southeast Asia region (Maldives), one country in the Americas region (Haiti), one country in the Eastern Mediterranean region (Afghanistan), three countries in the Europe region (Armenia, Azerbaijan, and Kyrgyz Republic), and both countries in the Western Pacific region (Philippines and Cambodia). Overall, about half of all countries demonstrating a three-class solution also showed some evidence for either a two- or four-class solution based on either CAIC or Likelihood-Ratio Test (LRT). In the subset of countries with strong or ambiguous evidence for a four-class model, no theoretically consistent class pattern emerged. WHO data also favored a three-class model for psychological violence variables only (100%). For the two countries where the findings for the LRT test suggested a solution other than a three-class solution (Peru and Thailand), the preponderance of evidence across the fit statistics suggested a three-class solution. Findings are somewhat similar for the all variables model, with all but two countries (Thailand and Samoa) having strong justification for a three-class model. However, all six had CAIC indices suggestive of a four-class

solution.

In Table 2, three-class models showed a low-to-no IPV class comprising the majority of the sample (DHS: 89.1%; WHO: 87.2%), a moderate intensity class comprising under 10% of the sample (DHS: 7.7%; WHO: 8.2%), and a high-intensity class comprising less than 5% of the sample (DHS: 3.2%; WHO: 4.7%). The low-to-no exposure class showed good homogeneity (similarity of individuals within the class) and gooddiscrimination (distinguishability of the class from other classes) and was characterized by high conditional probabilities of never experiencing all three items. A probability of an item response close to 0 or 1 implies that the particular response could be determined with great certainty given class membership (Ulbricht et al., 2018). The moderate-intensity class showed the highest probability of selecting "sometimes" for all items, while the high-intensity class showed the highest probabilities of selecting "often" for all items. In the moderate- and high-intensity classes, poor homogeneity (conditional probabilities between 0.3 and 0.7) was evident (Table 3), indicating suboptimal similarity in the responses for individuals within these classes. In addition, 20% of all odds ratios in these models fell between 0.2 and 5.0, considered suboptimal for class discrimination. Like poor homogeneity, poor discrimination was concentrated primarily in the moderate-versus high-intensity comparisons, especially among the odds ratios selecting never versus sometimes response options. This indicates that a response of never vs. sometimes to a given item did not adequately distinguish the moderate- and high-intensity classes.

The addition of physical and sexual IPV may have affected class structure, with greater ambiguity in the number of classes, although the three-class model still predominated. The addition of physical and sexual IPV did improve the discrimination of psychological violence items between the moderate- and high-intensity classes, with reductions from 42% to 27% and 44%–33% in the DHS and WHO data, respectively, in non-discriminating odds ratios. However, the addition of physical and sexual IPV did not improve homogeneity of the psychological violence items. Further, the physical and sexual IPV items themselves were very frequently implicated in poor heterogeneity and discrimination. Given the poor discrimination and heterogeneity of the physical and sexual IPV items, the lack of improvement in homogeneity, the greater ambiguity regarding the number of classes, and the added complexity of operationalizing the classes with the physical and sexual IPV items included, we chose to retain the three-class model based only on the psychological abuse items for further analysis.

3.3. Generalizability testing of the latent classes

For testing the similarity of the classes (Table 4), we included all countries (DHS = 47, WHO = 6) that demonstrated a three-class model

Table 1
Percentage of two-, three-, and four-class models in DHS and WHO data.

	Psychological IPV Items Only				Psychological IPV Items + Physical IPV + Sexual IPV			
	N	2 Classes	3 Classes	4 Classes	N	2 Classes	3 Classes	4 Classes
DHS								
Africa	27	0	27	0	24	3	18	3
Americas	4	0	4	0	4	0	3	1
E. Mediterranean	4	0	4	0	4	0	4	0
Europe	5	0	5	0	5	0	5	0
Southeast Asia	5	0	5	0	5	0	4	1
Western Pacific	2	0	2	0	2	0	2	0
Total	47	0	47	0	44	3	36	5
		0%	100%	0%		7%	82%	11%
WHO								
Africa	1	0	1	0	1	0	1	0
Americas	2	0	2	0	2	0	2	0
Southeast Asia	2	0	2	0	2	0	2	0
Western Pacific	1	0	1	0	1	0	1	0
Total	6	0	6	0	6	0	6	0
		0%	100%	0%		0%	100%	0%

Table 2

Conditional probabilities for three-class model, psychological violence only and all variables, 47 DHS countries[a] and 6 WHO countries.

Psychological IPV Items Only

DHS	Low to No IPV				Moderate Intensity IPV				High intensity IPV			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Class Prevalence	89.10%	5.60%	**67.50%**	97.50%	7.70%	4.10%	1.00%	23.50%	3.20%	1.90%	0.50%	9.00%
Insult: Never	0.94	0.03	0.85	0.99	0.15	0.09	0.04	**0.35**	0.08	0.08	0	**0.34**
Insult: Sometimes	0.05	0.03	0	0.13	0.75	0.1	**0.38**	0.92	0.07	0.06	0	0.2
Insult: Often	0.01	0.01	0	0.03	0.1	0.06	0	0.26	0.86	0.11	**0.5**	1
Humiliate: Never	0.98	0.01	0.95	1	0.29	0.15	0	0.82	0.17	0.12	0	**0.62**
Humiliate: Sometimes	0.02	0.01	0	0.04	**0.64**	0.14	0.17	0.96	0.09	0.06	0	0.23
Humiliate: Often	0	0	0	0.02	0.07	0.06	0.01	**0.37**	0.74	0.14	0.28	0.96
Threat: Never	0.99	0.01	0.96	1	**0.60**	0.15	0.18	0.88	**0.35**	0.16	0.02	0.8
Threat: Sometimes	0.01	0.01	0	0.03	**0.38**	0.15	0.10	0.78	0.14	0.11	0	**0.68**
Threat: Often	0	0	0	0.01	0.02	0.02	0	0.11	**0.51**	0.16	0.16	0.98

WHO	Low to No IPV				Moderate Intensity IPV				High intensity IPV			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Class Prevalence	87.20%	5.00%	79.30%	94.40%	8.20%	3.90%	2.90%	14.40%	4.70%	1.50%	2.70%	6.40%
Insult: Never	0.93	0.05	0.84	0.97	0.16	0.06	0.08	0.23	0.11	0.09	0.01	0.21
Insult: Sometimes	0.06	0.05	0.02	0.15	0.73	0.12	**0.6**	0.89	0.05	0.04	0	0.11
Insult: Often	0.01	0	0.01	0.02	0.11	0.06	0.03	0.17	0.85	0.1	**0.69**	0.96
Humiliate: Never	0.99	0	0.99	1	**0.44**	0.16	0.16	**0.65**	0.25	0.13	0.06	**0.44**
Humiliate: Sometimes	0.01	0	0	0.01	**0.52**	0.16	**0.34**	0.81	0.08	0.03	0.03	0.11
Humiliate: Often	0	0	0	0	0.03	0.03	0.01	0.1	**0.67**	0.12	**0.53**	0.87
Scare: Never	0.97	0.02	0.94	0.99	**0.42**	0.09	**0.32**	**0.56**	0.25	0.06	0.19	**0.35**
Scare: Sometimes	0.03	0.01	0.01	0.05	**0.51**	0.1	**0.34**	**0.6**	0.13	0.03	0.09	0.17
Scare: Often	0.01	0	0	0.01	0.07	0.03	0.04	0.1	**0.62**	0.07	**0.53**	0.71
Threat: Never	0.99	0	0.99	1	**0.69**	0.11	**0.61**	0.91	**0.43**	0.19	0.26	0.74
Threat: Sometimes	0.01	0	0	0.01	0.29	0.11	0.08	**0.38**	0.17	0.08	0.06	0.26
Threat: Often	0	0	0	0	0.02	0.01	0	0.04	**0.4**	0.17	0.15	**0.66**

Psychological IPV Items + Physical IPV + Sexual IPV

DHS	Low to No IPV				Moderate Intensity IPV				High intensity IPV			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Class Prevalence	84.70%	6.50%	**66.00%**	95.90%	11.70%	5.00%	2.60%	22.30%	3.60%	2.10%	0.60%	11.80%
Insult: Never	0.97	0.02	0.93	1	0.3	0.19	0.02	0.88	0.09	0.08	0	**0.32**
Insult: Sometimes	0.03	0.02	0	0.07	**0.61**	0.18	0.1	0.88	0.12	0.1	0	**0.49**
Insult: Often	0.01	0.01	0	0.02	0.08	0.06	0	0.21	0.79	0.15	**0.31**	1
Humiliate: Never	0.98	0.01	0.96	1	**0.44**	0.16	0.09	0.78	0.18	0.1	0.04	**0.46**
Humiliate: Sometimes	0.01	0.01	0	0.03	**0.49**	0.15	0.19	0.88	0.13	0.1	0	**0.51**
Humiliate: Often	0	0	0	0.01	0.07	0.07	0	**0.31**	0.7	0.12	**0.43**	0.93
Threat: Never	1	0	0.99	1	0.73	0.13	0.28	0.97	**0.38**	0.18	0.04	0.8
Threat: Sometimes	0	0	0	0.01	0.25	0.12	0.03	**0.54**	0.18	0.13	0	**0.53**
Threat: Often	0	0	0	0.01	0.02	0.03	0	0.19	**0.44**	0.18	0.1	0.96
Sexual IPV: None	0.93	0.22	0	1	0.73	0.22	0	0.99	**0.56**	0.21	0	0.93
Sexual IPV: Any	0.07	0.22	0	1	0.27	0.22	0.01	1	**0.44**	0.21	0.08	1
Physical IPV: None	0.94	0.06	0.73	0.99	**0.35**	0.16	0.08	**0.69**	0.18	0.11	0	**0.49**
Physical IPV: Non-severe	0.04	0.03	0.01	0.15	0.27	0.09	0.13	**0.56**	0.13	0.09	0	**0.45**
Physical IPV: Severe	0.02	0.03	0	0.12	**0.38**	0.13	0.09	**0.68**	0.69	0.16	**0.36**	1

WHO	Low to No IPV				Moderate Intensity IPV				High intensity IPV			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Class Prevalence	82.10%	5.90%	72.20%	87.30%	13.30%	5.10%	8.40%	21.40%	4.60%	1.50%	2.70%	6.40%
Insult: Never	0.95	0.04	0.87	0.98	**0.31**	0.12	0.19	**0.51**	0.11	0.1	0	0.22
Insult: Sometimes	0.05	0.04	0.02	0.12	**0.58**	0.12	**0.41**	0.75	0.05	0.03	0	0.08
Insult: Often	0.01	0	0	0.01	0.11	0.04	0.04	0.16	0.84	0.1	0.7	0.95
Humiliate: Never	0.99	0.01	0.99	1	**0.6**	0.16	**0.35**	0.79	0.25	0.14	0.07	**0.47**
Humiliate: Sometimes	0.01	0	0	0.01	**0.37**	0.16	0.17	**0.63**	0.1	0.05	0.03	0.15
Humiliate: Often	0	0	0	0	0.03	0.01	0.01	0.04	**0.66**	0.13	**0.5**	0.87
Scare: Never	0.98	0.02	0.95	1	**0.54**	0.12	**0.34**	**0.68**	0.24	0.06	0.18	**0.35**
Scare: Sometimes	0.02	0.01	0	0.04	**0.4**	0.11	0.27	**0.56**	0.14	0.05	0.09	0.21
Scare: Often	0	0	0	0.01	0.07	0.03	0.03	0.1	**0.62**	0.08	**0.54**	0.71
Threat: Never	1	0	0.99	1	0.77	0.1	**0.67**	0.91	**0.41**	0.2	0.2	0.75
Threat: Sometimes	0	0	0	0.01	0.21	0.09	0.07	**0.31**	0.18	0.11	0.04	0.29
Threat: Often	0	0	0	0	0.02	0.01	0	0.03	**0.41**	0.18	0.15	**0.68**
Sexual IPV: None	0.93	0.05	0.86	0.99	**0.59**	0.14	**0.45**	0.85	**0.45**	0.08	**0.35**	**0.57**
Sexual IPV: Any	0.07	0.05	0.01	0.15	**0.42**	0.14	0.15	**0.55**	**0.55**	0.08	**0.43**	**0.65**
Physical IPV: None	0.95	0.03	0.91	0.98	**0.4**	0.09	0.25	**0.49**	0.26	0.12	0.08	**0.41**
Physical IPV: Non-severe	0.03	0.02	0.01	0.06	0.26	0.06	0.19	**0.35**	0.12	0.05	0.06	0.19
Physical IPV: Severe	0.02	0.01	0	0.04	**0.35**	0.14	0.16	**0.57**	0.62	0.14	**0.48**	0.86

[a] Conditional probability ranges between 0.300 and 0.700 are in bold.

Table 3

Number and percentage of countries with non-discriminating odds ratios and non-homogeneous conditional probabilities in three-class models, psychological only and all IPV items, DHS and WHO data.

Non-Discriminating Odds Ratios (between 0.2 and 5.0)

DHS	Psychological IPV Items Only						Psychological IPV Items + Physical IPV + Sexual IPV					
	Low vs. Moderate	%	Low vs. High	%	Moderate vs. High	%	Low vs. Moderate	%	Low vs. High	%	Moderate vs. High	%
Insult never vs. sometimes	3	6%	4	9%	33	70%	0	0%	4	9%	16	36%
Insult sometimes vs. often	1	2%	8	17%	3	6%	6	14%	6	14%	0	0%
Humiliate never vs. sometimes	0	0%	0	0%	42	89%	0	0%	1	2%	26	59%
Humiliate sometimes vs. often	4	9%	5	11%	4	9%	4	9%	5	11%	3	7%
Threat never vs. sometimes	0	0%	0	0%	35	74%	5	11%	5	11%	22	50%
Threat sometimes vs. often	2	4%	6	13%	2	4%	2	5%	8	18%	4	9%
Sexual IPV any vs. none							4	9%	3	7%	35	80%
Physical IPV moderate vs. none							0	0%	2	5%	34	77%
Physical IPV high vs. moderate							0	0%	1	2%	27	61%
Average %		4%		8%		42%		5%		9%		42%
Average psychological violence %								6%		11%		27%

WHO	Psychological IPV Items Only						Psychological IPV Items + Physical IPV + Sexual IPV					
	Low vs. Moderate	%	Low vs. High	%	Moderate vs. High	%	Low vs. Moderate	%	Low vs. High	%	Moderate vs. High	%
Insult never vs. sometimes	0	0%	0	0%	4	67%	0	0%	0	0%	5	83%
Insult sometimes vs. often	1	17%	0	0%	0	0%	0	0%	0	0%	0	0%
Humiliate never vs. sometimes	0	0%	0	0%	5	83%	0	0%	0	0%	3	50%
Humiliate sometimes vs. often	1	17%	0	0%	0	0%	0	0%	1	17%	0	0%
Scare never vs. sometimes	0	0%	0	0%	6	100%	0	0%	0	0%	4	67%
Scare sometimes vs. often	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
Threat never vs. sometimes	0	0%	0	0%	6	100%	0	0%	0	0%	3	50%
Threat sometimes vs. often	0	0%	0	0%	0	0%	1	17%	2	33%	1	17%
Physical IPV moderate vs. none							0	0%	0	0%	6	100%
Physical IPV high vs. moderate							0	0%	0	0%	5	83%
Sexual IPV vs. none							0	0%	0	0%	6	100%
Average %		4%		0%		44%		2%		5%		50%
Average psychological violence %								2%		6%		33%

Poor Homogeneity (Conditional Probabilities between 0.3 and 0.7)

DHS	Psychological IPV Items Only						Psychological IPV Items + Physical IPV + Sexual IPV					
	Low	%	Mod	%	High	%	Low	%	Mod	%	High	%
Insult never	0	0%	4	9%	2	4%	0%	0%	19	43%	1	2%
Insult sometimes	0	0%	14	30%	0	0%	0%	0%	26	59%	2	5%
Insult often	0	0%	1	2%	5	11%	0%	0%	0	0%	9	20%
Humiliate never	0	0%	20	43%	3	6%	0%	0%	33	75%	6	14%
Humiliate sometimes	0	0%	31	66%	1	2%	0%	0%	35	80%	3	7%
Humiliate often	0	0%	1	2%	14	30%	0%	0%	1	2%	22	50%
Threat never	0	0%	35	74%	29	62%	0%	0%	16	36%	26	59%
Threat sometimes	0	0%	32	68%	3	6%	0%	0%	13	30%	5	11%
Threat often	0	0%	0	0%	38	81%	0%	0%	0	0%	33	75%
Sexual IPV never							0%	0%	11	25%	25	57%
Sexual IPV any							0%	0%	11	25%	25	57%
Physical IPV never							0%	0%	27	61%	6	14%
Physical IPV moderate							0%	0%	16	36%	2	5%
Physical IPV High							0%	0%	32	73%	20	45%
Average %		**0%**		**33%**		**22%**		**0%**		**39%**		**30%**
Average psychological violence %								**0%**		**36%**		**27%**

Table 3 (*continued*)

Poor Homogeneity (Conditional Probabilities between 0.3 and 0.7)												
DHS	Psychological IPV Items Only						Psychological IPV Items + Physical IPV + Sexual IPV					
	Low	%	Mod	%	High	%	Low	%	Mod	%	High	%
WHO	**Psychological violence Only**						**Psychical IPV Items + Physical IPV + Sexual IPV**					
	Low	**%**	**Mod**	**%**	**High**	**%**	**Low**	**%**	**Mod**	**%**	**High**	**%**
Insult never	0	0%	0	0%	0	0%	0	0%	2	33%	0	0%
Insult sometimes	0	0%	3	50%	0	0%	0	0%	4	67%	1	17%
Insult often	0	0%	0	0%	1	17%	0	0%	0	0%	0	0%
Humiliate never	0	0%	5	83%	2	33%	0	0%	4	67%	2	33%
Humiliate sometimes	0	0%	5	83%	0	0%	0	0%	4	67%	0	0%
Humiliate often	0	0%	0	0%	4	67%	0	0%	1	17%	4	67%
Scare never	0	0%	6	100%	1	17%	0	0%	5	83%	2	33%
Scare sometimes	0	0%	6	100%	0	0%	0	0%	4	67%	1	17%
Scare often	0	0%	0	0%	4	67%	0	0%	1	17%	4	67%
Threat never	0	0%	4	67%	3	50%	0	0%	3	50%	2	33%
Threat sometimes	0	0%	4	67%	0	0%	0	0%	1	17%	0	0%
Threat often	0	0%	0	0%	5	83%	0	0%	1	17%	4	67%
Physical IPV never							0	0%	4	67%	3	50%
Physical IPV moderate							0	0%	1	17%	0	0%
Physical IPV High							0	0%	5	83%	4	67%
Sexual IPV never							0	0%	5	83%	6	100%
Sexual IPV any							0	0%	5	83%	6	100%
Average %		**0%**		**50%**		**22%**		**0%**		**49%**		**38%**
Average psychological violence %								**0%**		**42%**		**28%**

in the exploratory LCA analyses of the psychological abuse items only. Because Africa was much larger than all other regions at 27 countries, we split the region into three geographic subregions (West, East, and Southern/Central). Based on improvement of fit statistics (namely a reduction in the CAIC from the less to the more constrained model), we found the item response probabilities, conditional on class, to be the same for all countries within regions, except for the Eastern Mediterranean region. After constraining the class prevalences to be the same,

Table 4
Results of regional and cross-regional generalizability testing.

Region	Countries	N	k	LL	#fp	AIC	BIC	CAIC	SABIC	Entropy	LR Test	p-value
West Africa	12											
Configural		49348	3	−134470.16	251	269442.32	271652.79	270369.33	270855.11	0.956		
Structural		49348	3	−124422.19	53	248950.39	249417.14	249146.13	249248.71	0.962	5375.73	<0.001
Distributional		49348	3	−124838.27	31	249738.53	250011.54	249853.02	249913.02	0.965	504.11	<0.001
East Africa	6											
Configural		33635	3	−88555.62	125	177361.24	178414.15	177802.08	178016.90	0.955		
Structural		33635	3	−88774.66	41	177631.31	177976.67	177775.91	177846.37	0.952	233.49	<0.001
Distributional		33635	3	−89183.53	25	178417.05	178627.63	178505.22	178548.18	0.955	327.42	<0.001
Southern & Central Africa	9											
Configural		43248	3	−121215.01	188	242806.01	244436.86	243489.57	243839.39	0.959		
Threshold		43248	3	−121616.13	44	243320.26	243701.94	243480.24	243562.11	0.951	578.37	<0.001
Distributional		43248	3	−122467.56	28	244991.12	245234.01	245092.93	245145.03	0.965	577.85	<0.001
Americas	4											
Configural		29122	3	−66268.67	83	132703.33	133390.51	132990.86	133126.74	0.966	a	
Structural		29122	3	−66372.34	29	132802.68	133042.77	132903.14	132950.61	0.967	a	
Distributional		29122	3	−66606.29	23	133258.58	133449.01	133338.26	133375.91	0.965	a	
Eastern Mediterranean	4											
Configural		38717	3	−88506.34	83	177178.69	177889.50	177476.48	177625.73	0.942		
Structural		38717	3	−90079.27	29	180216.55	180464.91	180320.60	180372.74	0.914	1140.58	<0.001
Distributional		–	–	–	–	–	–	–	–	–	–	–
Europe	5											
Configural		20392	3	−38925.30	104	78058.60	78882.58	78402.78	78552.07	0.956		
Structural		20392	3	−39076.62	32	78217.25	78470.78	78323.15	78369.09	0.978	237.82	<0.001
Distributional		20392	3	−39379.49	24	78806.97	78997.12	78886.40	78920.85	0.978	−2141.13	**
Southeast Asia	5											
Configural		80346	3	−73860.03	104	147928.06	148894.65	148334.18	148564.13	0.975		
Structural		80346	3	−74022.93	32	148109.86	148407.27	148234.82	148305.57	0.979	444.50	<0.001
Distributional		80346	3	−73959.46	24	147966.92	148189.97	148060.64	148113.70	0.979	−219.47	**
Western Pacific	2											
Configural		16712	3	−14678.55	41	29438.89	29755.67	29571.24	29625.27	0.925		
Threshold		16712	3	−14716.26	23	29478.52	29656.17	29552.65	29583.08	0.973	69.38	<0.001
Distributional		16712	3	−15105.11	21	30252.22	30414.42	30319.90	30347.68	0.973	963.27	<0.001
All Regions (without EasternMediterranean)	43											
Configural		266960	3	−563897.69	146	1128087.37	1129619.62	1128733.64	1129155.63	0.971		
Threshold		266960	3	−570495.49	38	1141066.98	1141465.79	1141235.19	1141345.02	0.923	6036.05	<0.001
Distributional		–	–	–	–	–	–	–	–	–	–	–

**P-value could not be calculated due to a negative value for scaling correction.

[a] LRT could not be calculated due to an issue with the information matrix.

we found that this constraint held only in the southeast Asia region, suggesting similar class prevalences across the five countries in this region. In the pooled analysis of all 8 DHS regions and subregions demonstrating cross-country invariance in conditional probabilities, we found that the response category probabilities were not invariant, suggesting regional differences in the probability of choosing at or above a response category of an item given latent class membership. The WHO countries were too few to fully replicate the regional equivalence testing, but a multi-country test of the WHO data confirmed the DHS findings of a lack of equivalence in the conditional item response probabilities (results not shown).

3.4. Class operationalization and validity testing

Given the consistency of the three-class model in individual country and region-specific models, but the lack of global similarity, we developed a classification strategy that was general enough to accommodate potential differences in response probabilities of specific variables, but also consistent with the nature of the three-class model. For the DHS, we classified individuals in the operationalized moderate-intensity class if they either responded "sometimes" to both insult and humiliate or if they responded "sometimes" to threat. Individuals who responded "often" to both insult and humiliate, or who responded "often" to threat, were classified in the high-intensity class. Given the additional item "scare" in the WHO dataset, responding "sometimes" to at least two of insult, humiliate/belittle, and scare or responding "sometimes" to threat alone was the criterion for classification as moderate-vs. low-to-none. Similarly, responding "often" to at least two of insult, humiliate/belittle,

and scare or responding "often" to threat alone was the criterion for classification as high-intensity. This classification strategy closely replicates the distribution of the latent three-class model, with roughly 90% falling in the low class, 7% in the moderate class, and 3% in the severe class in both the DHS and WHO datasets. In models using operationalized classes, the latent classes showed a dose response relationship with injury due to IPV, physical and sexual IPV and controlling behaviors (Fig. 1a and b). The pattern of higher relative risk among those in the high-intensity class was similar in the WHO data, although the confidence intervals more frequently overlapped for countries in Southeast Asia and the Western Pacific.

3.5. Item response theory analysis

The three DHS and four WHO items captured the latent construct of psychological IPV well between 0.5 and 3.0 standard deviations above the mean. This was consistent across countries and datasets. High information coverage across a broad range of the latent trait reflects good coverage of a latent construct. In general, item discriminations were high (>2). For most countries, item information curves overlapped considerably (Fig. 2), suggesting that the items were measuring roughly the same level of psychological violence intensity with a similar level of precision. Overall, this may help to explain the lack of clearer class separations for the moderate and high groups shown in Table 3. In both the DHS and WHO datasets, there was substantial overlap in the information provided by the "insult" and "humiliate" items. "Insult" provided the most information in the majority of DHS countries (24/47) and regions, followed by "humiliate" (18/47 countries), with some variation
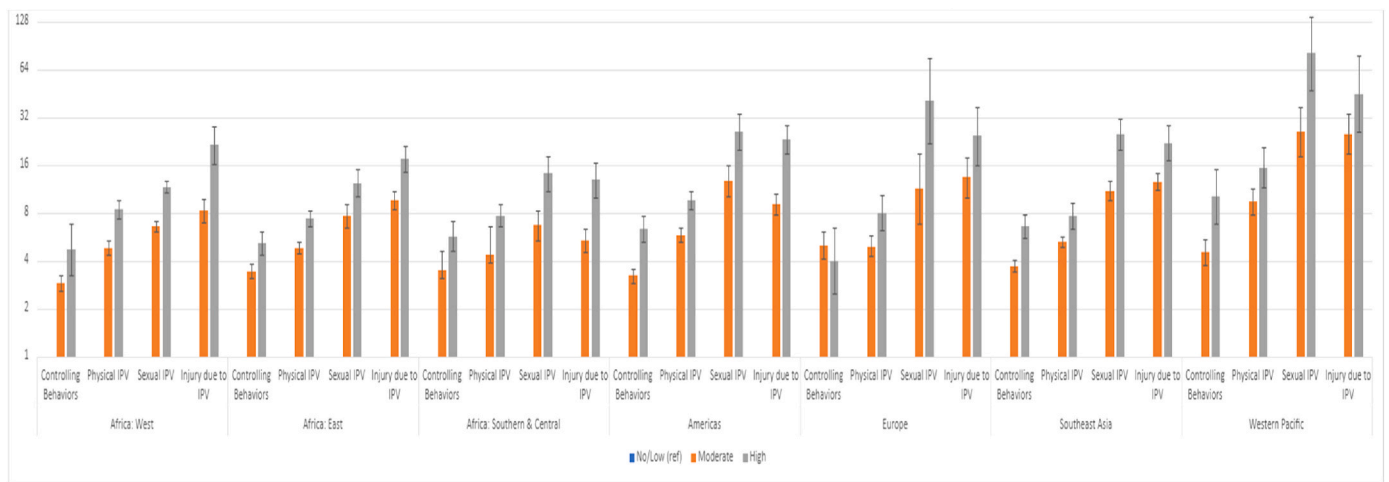

Fig. 1a. Relative Risk Ratios of Class Membership by related outcomes, DHS by Region.
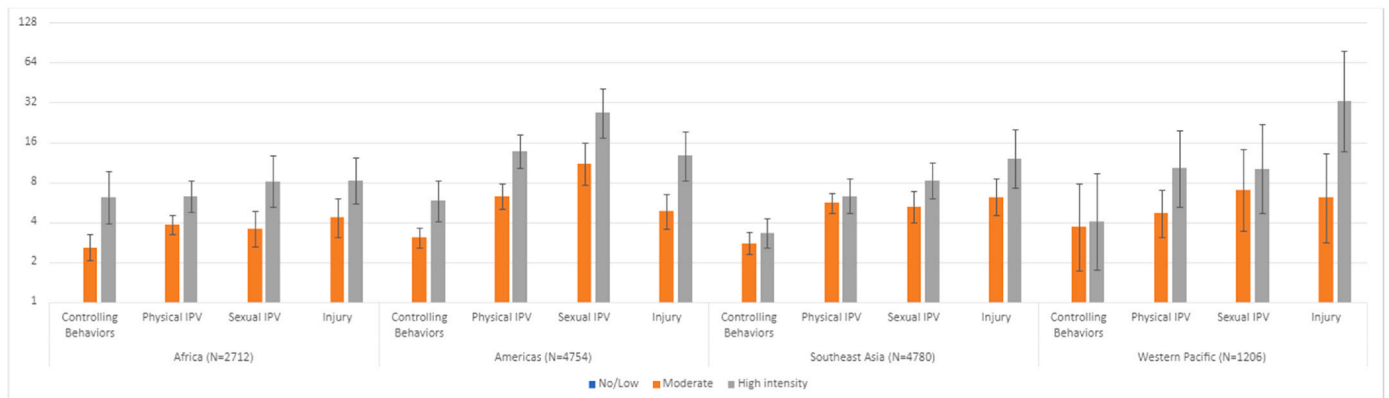

Fig. 1b. Relative risk ratios of class membership by other IPV exposures, WHO by country.
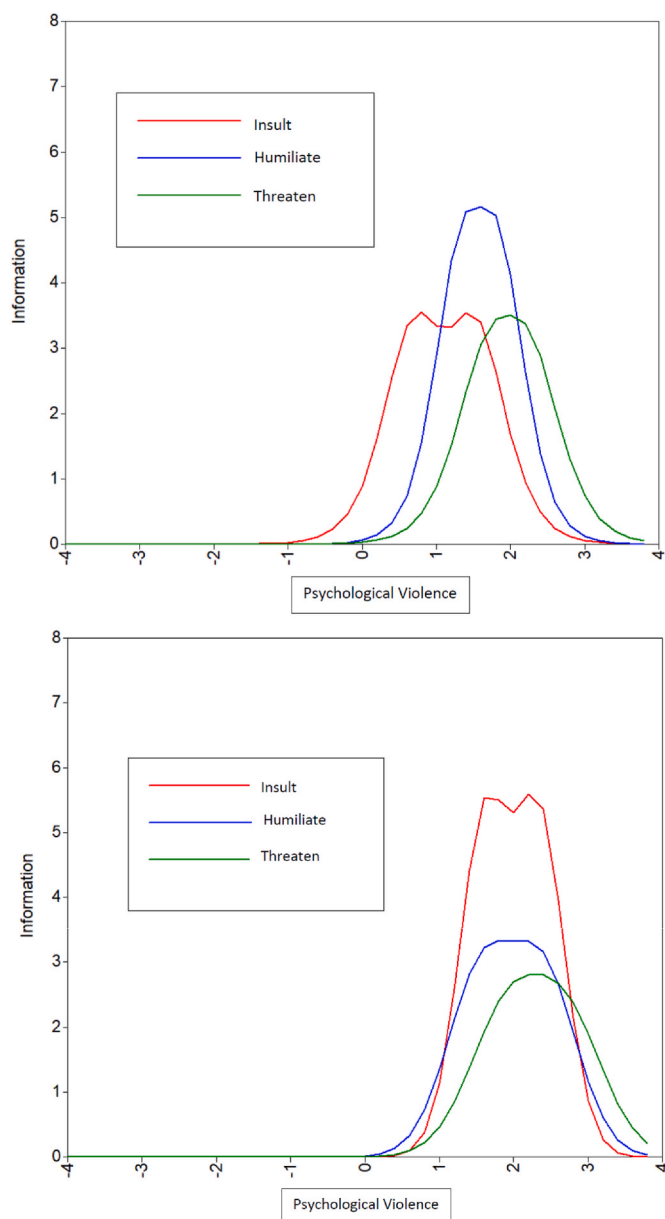
Fig. 2. Item information curves, Tanzania (top), India (bottom), DHS.

by region; for example, "humiliate" provided the most information in all four Eastern Mediterranean countries. In the WHO data, the "scare/intimidate" item covered the same range of psychological IPV as humiliate while contributing less information. "Threat" had higher difficulty than the other items in both datasets, reflecting more severe psychological violence. Overall, the items used in both the DHS and the WHO datasets are measuring a range of severity of psychological IPV, althoughthere is considerable overlap in the content measured by some items across a number of datasets, especially for the "scare/intimidate" and "humiliate" items.

4. Discussion

4.1. Statement of principal findings

This is the largest analysis of the patterning and item characteristics of women's experiences of psychological IPV against women in LMICs to date. The three-class model found in smaller-scale studies (Clark et al., 2019) predominated (high-intensity, moderate-intensity, low-to-no

exposure), and was confirmed by regional analyses. While the consistency and within-region structural similarity of the three-class model is encouraging, this study also found a lack of homogeneity and poor discrimination in the moderate- and high-intensity classes. Further, when pooled at the global level, the class structure was variant, suggesting some regional differences in the exact nature of the classes. Given the general consistency of the three-class solution, but also some deficiencies in homogeneity and discrimination of the classes and the lack of global measurement equivalence, the team proposes operationalized classes that most closely represent the fundamental structure of the LCA three-class solution while also generalized enough to accommodate the differences detectable in the pooled global sample. Consistently, these operationalized classes demonstrate predictive validity via a graded relationship with other IPV-related outcomes. In addition, IRT analysis confirmed that the three DHS and four WHO items captured the latent construct of psychological IPV well between 0.5 and 3.0 standard deviations above the mean with moderate to high item discriminations, but some redundancy among the items in terms of the severity of violence measured. Overall, this suggests that three-four items are sufficient to capture a broad range of women's experience of IPV and to distinguish classes of women experiencing various levels of psychological violence, and that refinement of the item set could further improve measurement of psychological IPV among women in LMICs.

4.2. Strengths and weaknesses in relation to other studies

The analysis is the only comprehensive assessment of the patterning of women's experience of psychological IPV including a confirmatory approach and testing for the generalizability of latent profiles, a prerequisite for cross-country use and comparisons. The study confirmed a 3-class model identified in prior exploratory testing on the WHO sample (Heise et al., 2019a), and studies in Ireland (Watson & Parsons, 2005) and the US(Saint-Eloi Cadely et al., 2020) in which psychological violence was modeled separately from other forms of IPV. Also consistent with prior latent class research on IPV that either modeled psychological violence separately (Watson & Parsons, 2005) or in combination with other forms of IPV(Dutton et al., 2005; Haynie et al., 2013; McNaughton Reyes et al., 2018; Restrepo et al., 2022; Spencer et al., 2016; Weiss et al., 2017) is a graded class structure and the presence of a high intensity class. This class, while small, is associated with the greatest risk of other forms of IPV, supporting prior work identifying an overlap in forms of IPV and elevated risk of range of adverse social and health outcomes for individuals experiencing high intensity or severe violence (Ansara & Hindin, 2011; Carbone-López et al., 2006; Clark et al., 2019; Dutton et al., 2005; Haynie et al., 2013; Heise et al., 2019b; McNaughton Reyes et al., 2021; Weiss et al., 2017). However, this analysis also demonstrated the difficulties of clearly separating this class from the moderate intensity class, both of which are clearly distinguishable from the no-to-low class, but not definitively distinguishable from one other. This study's unique modeling of frequency, which is a design feature in only one prior study, the exploratory study on WHO data, identified culprits to the lack of class separation, namely the lowest frequency items, which across both datasets, contributed to poor discrimination.

This analysis differs from the exploratory WHO analysis in its examination of the impact of simultaneously including physical and sexual IPV in the models on the patterning of psychological IPV. Given the overlap in women's experience of different forms of IPV shown in prior research (World Health Organization, 2005), we anticipated that the patterning of women's experience of psychological violence would differ when these other experiences were accounted for. The 3-class model remained the most frequently identified, like in a recent analysis of DHS data from 5 Latin American countries (Restrepo et al., 2022), and the addition of physical and sexual IPV increased the discrimination of the moderate-versus high-intensity classes. However, homogeneity did not improve, the structure became more variable and less

interpretable while the operationalization of the classes became more complex. The lack of consistent overall benefit to the classification of women's experience of psychological violence according to their concurrent experience of physical and sexual IPV may be due to considerable variability in the patterning of women's experiences of physical and sexual IPV which themselves demonstrated poor homogeneity and discrimination in the models.

4.3. Limitations and strengths

Firstly, our sample consisted exclusively of low- and middle-income countries (LMICs) and patterning of women's experience of IPV in high-income countries may differ from those in LMICs. However, the limited prior research examining patterns in women's experience of psychological violence (without simultaneously modeling other forms of violence, risk factors, or outcomes) has occurred in high income settings and identified a similar 3-class solution, suggesting that the pattern might not differ. Although WHO and DHS items were broadly similar, they differ in number and the nomenclature of the frequencies. However, the proposed operationalization takes into consideration this limitation. The inclusion of all ever-partnered women in the analytic sample also result in underestimates of IPV prevalence, as women who experienced past IPV only prior to the past 12 months would be coded as not experiencing abuse. Furthermore, the WHO data include only one urban area and one rural area in most of the included sites, which is less representative than the DHS. The broad similarities in the 3-class solution across the two datasets and with existing research suggests that this limitation is modest at best. Caution is warranted due to the distribution of responses to psychological IPV items, as data with a preponderance of zeroes may produce false negatives in LCA. Replication of the findings with "ever" violence or in samples with higher prevalence is needed to ascertain the degree to which the patterns established with low prevalence items are impacted by the low positive endorsement of the items. Furthermore, the amount of information is limited by the small number of items (Wurpts & Geiser, 2014). Finally, the suggested operationalization requires further testing and is based on imperfect data which showed very promising, but not global patterning. Despite these limitations, our findings are strengthened by large samples across two data sets and a methodologically rigorous approach to testing measurement equivalence.

4.4. Implications for research

As has been shown through the Millennium Development Goals process, the establishment of specific indicators leads to progress in their measurement (United Nations, 2015a) the same is expected of the measurement of IPV through the SDG process. The regularity and standardization of data collection has expanded, but questions remain regarding measurement and the meaning and boundaries of the construct across countries and cultures. This analysis has shown, across nearly 60 LMICs, that a three-class model is conceptually and structurally similar across countries within regions. However, the lack of full measurement equivalence in the pooled global sample suggests that additional research including in-depth cognitive testing, and cross-national qualitative studies and survey experiments are needed to clarify the set of items that unequivocally demarcate low-, moderate- and high-intensity psychological violence in a globally comparable way.

First, cross-context mixed methods research is needed to establish common conceptual boundaries of psychological violence to rectify the persistent lack of consensus on the definition and measurement of psychological IPV. The non-zero conditional probabilities of experiencing psychological and other forms of IPV among the otherwise "no" violence class, and the contribution of the lower frequency items to poor homogeneity and demarcation of the classes suggests that not all items and not all frequencies are indicative of violence and its severity. It also suggests that the common practice of conceptualizing any positive response to

any item of any frequency is overestimating the prevalence of psychological IPV and obfuscating the much greater risk of social and health consequences experienced by women in the high intensity class.

Second, the number of items available for testing is small, only three or four in the most frequently used tools to monitor SDGs. In the present study, no class structure exceeded three in the DHS, which has only three items. In the WHO, there was the suggestion of four classes, a possibility that increased with the addition of the physical and sexual IPV items. Given the need for short scales for multi-purpose surveys like the DHS or even violence surveys that assess multiple forms of violence, it is unlikely that large scales are feasible. However, it is highly unlikely that a construct as complex as psychological violence (Follingstad & Rogers, 2013) can claim content validity with only three or four items, particularly with the large degree of overlap observed in item information curves. The addition of items that measure a wider range of severity would enhance the information obtained through the scale. Taking an inventory of theory and existing validated item sets used to measure psychological violence may inform future additions to WHO and DHS items to expand the domain of the underlying construct that is captured.

Third, a quantitative classification of frequency is needed to remove some of the variability in interpretation that the current subjective frequencies seem to be eliciting. Finally, some variability in the class structure could be due to differences in survey administration or quality of translation, reaffirming the need for violence-related surveys to consistently follow best practices in these domains. Ultimately, the proposed cross-context research agenda would aim to clarify the definition of psychological violence, provide a comprehensive understanding of its domains and lead to a core item set that is globally comparable across low- and middle-income and high income countries and regions, item sub-sets that are context specific, and ultimately consensus on best practices in the quantitative measurement, reporting and modeling of prevalence estimates of psychological IPV.

5. Conclusion

This study offers novel confirmation of a broadly comparable pattern of women's experience of psychological violence in LMICs, and in doing so, provides insight into the demarcation of abuse compared to other deleterious experiences while also providing severity cut points worthy of future validation tests. Our study also highlights key challenges in the measurement of psychological violence, including limited item sets, inconsistent item wording and variations in frequency measures across studies. Through our findings, we aim to facilitate further refinement of psychological violence measures for global monitoring including for tracking progress on SDG 5.2.1.

Funding

CRediT authorship contribution statement

Cari Jo Clark: Conceptualization, Formal analysis, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Irina Bergenfeld:** Conceptualization, Data curation, Formal analysis, Methodology,

Resources, Visualization, Writing – original draft, Writing – review & editing. **Yuk Fai Cheong:** Conceptualization, Formal analysis, Methodology, Writing – review & editing. **Hector Najera:** Formal analysis, Methodology, Writing – review & editing. **LynnMarie Sardinha:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing. **Claudia García-Moreno:** Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. **Lori Heise:** Conceptualization, Formal analysis, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

DHS datasets are available for download by submitting a request at https://dhsprogram.com/data/available-datasets.cfm. The WHO data are not publicly available.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssmph.2023.101500.

References

Ansara, D. L., & Hindin, M. J. (2011). Psychosocial consequences of intimate partner violence for women and men in Canada. *Journal of Interpersonal Violence, 26*(8), 1628–1645. https://doi.org/10.1177/0886260510370600

Bichi, A. A., & Talib, R. (2018). Item response theory: An introduction to latent trait models to test and item development. *International Journal of Evaluation and Research in Education, 7*(2), 142–151.

Carbone-López, K., Kruttschnitt, C., & Macmillan, R. (2006). Patterns of intimate partner violence and their associations with physical health, psychological distress, and substance use. *Public Health Reports, 121*(4), 382–392. http://www.ncbi.nlm.nih.gov/pubmed/16827439.

Clark, C. J., Cheong, Y. F., Gupta, J., Ferguson, G., Shrestha, B., Shrestha, P. N., & Yount, K. M. (2019). Intimate partner violence in Nepal: Latent patterns and association with depressive symptoms. *SSM Popul Health, 9*, Article 100481. https://doi.org/10.1016/j.ssmph.2019.100481

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences.* John Wiles & Sons.

Devries, K. M., Mak, J. Y., Garcia-Moreno, C., Petzold, M., Child, J. C., Falder, G., Lim, S., Bacchus, L. J., Engell, R. E., & Rosenfeld, L. (2013). The global prevalence of intimate partner violence against women. *Science, 340*(6140), 1527–1528.

Dutton, M. A., Kaltman, S., Goodman, L. A., Weinfurt, K., & Vankos, N. (2005). Patterns of intimate partner violence: Correlates and outcomes. *Violence & Victims, 20*(5), 483–497. https://www.ncbi.nlm.nih.gov/pubmed/16248486.

Finch, H. (2015). A comparison of statistics for assessing model invariance in latent class analysis. *Open Journal of Statistics, 5*, 191–210.

Follingstad, D. R., & Rogers, M. J. (2013). Validity concerns in the measurement of women's and men's report of intimate partner violence. *Sex Roles, 69*(3–4), 149–167.

García-Moreno, Claudia., Jansen, H. A. F. M., Ellsberg, M., Heise, L., & Watts, C. (2005). *WHO Multi-country Study on Women's Health and Domestic Violence against Women: Initial results on prevalence, health outcomes and women's responses W. H. Organization.*

Garcia-Moreno, C., Jansen, H. A., Ellsberg, M., Heise, L., & Watts, C. H. (2006). Prevalence of intimate partner violence: Findings from the WHO multi-country study on women's health and domestic violence. *Lancet, 368*(9543), 1260–1269. https://doi.org/10.1016/S0140-6736(06)69523-8. S0140-6736(06)69523-8 [pii].

Halim, N., Beard, J., Mesic, A., Patel, A., Henderson, D., & Hibberd, P. (2018). Intimate partner violence during pregnancy and perinatal mental disorders in low and lower middle income countries: A systematic review of literature. *1990–2017. Clinical psychology review, 66*, 117–135.

Haynie, D. L., Farhat, T., Brooks-Russell, A., Wang, J., Barbieri, B., & Iannotti, R. J. (2013). Dating violence perpetration and victimization among U.S. Adolescents: Prevalence, patterns, and associations with health complaints and substance use.

Journal of Adolescent Health, 53(2), 194–201. https://doi.org/10.1016/j.jadohealth.2013.02.008

Health, W. D. o. o G. W. a.. (2001). *Putting women first: Ethical and safety recommendations for research on domestic violence against women.* W. H. Organization.

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

Heise, L., Pallitto, C., García-Moreno, C., & Clark, C. J. (2019b). Measuring psychological abuse by intimate partners: Constructing a cross-cultural indicator for the Sustainable Development Goals. *SSM-Population Health, 9*, Article 100377.

Martín-Fernández, 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), 1–11.

McNaughton Reyes, H. L., Foshee, V. A., Chen, M. S., Gottfredson, N. C., & Ennett, S. T. (2018). Consequences of involvement in distinct patterns of adolescent peer and dating violence. *Journal of Youth and Adolescence, 47*, 2371–2383.

McNaughton Reyes, H. L., Maman, S., Chen, M. S., Groves, A. K., & Moodley, D. (2021). Patterns of intimate partner violence victimization among South African women and their relation to emotional distress during pregnancy and postpartum. *Journal of Interpersonal Violence, 36*(7–8), NP4230–NP4249.

Muthén, L., & Muthén, B. (2018). *Mplus user's guide.* Version, 2017 8. .

Perra, O. (2020). *Latent transition analysis.* SAGE Publications Limited.

Restrepo, A., Montoya, N., & Zuluaga, L. (2022). Typologies of intimate partner violence against women in five Latin-American countries: A latent class analysis. *International Journal of Public Health, 150.*

Saint-Eloi Cadely, H., Pittman, J. F., Pettit, G. S., Lansford, J. E., Bates, J. E., Dodge, K. A., & Holtzworth-Munroe, A. (2020). Classes of intimate partner violence from late adolescence to young adulthood. *Journal of Interpersonal Violence, 35*(21–22), 4419–4443.

Salis, K. L., Salwen, J., & O'Leary, K. D. (2014). The predictive utility of psychological aggression for intimate partner violence. *Partner Abuse, 5*(1), 83–97.

Sardinha, L., Maheu-Giroux, M., Stöckl, H., Meyer, S. R., & García-Moreno, C. (2022). Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. *The Lancet, 399*(10327), 803–813.

Schmiege, S. J., Masyn, K. E., & Bryan, A. D. (2018). Confirmatory latent class analysis: Illustrations of empirically driven and theoretically driven model constraints. *Organizational Research Methods, 21*(4), 983–1001.

Semahegn, A., Torpey, K., Manu, A., Assefa, N., Tesfaye, G., & Ankomah, A. (2019). Are interventions focused on gender-norms effective in preventing domestic violence against women in low and lower-middle income countries? A systematic review and meta-analysis. *Reproductive Health, 16*(1), 93. https://doi.org/10.1186/s12978-019-0726-5

Spencer, R. A., Renner, L. M., & Clark, C. J. (2016). Patterns of dating violence perpetration and victimization in US young adult males and females. *Journal of Interpersonal Violence, 31*(15), 2576–2597.

StataCorp, L. (2019). *Stata statistical software: Release 16. College station, TX.*

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.

Toland, M. D. (2014). Practical guide to conducting an item response theory analysis. *The Journal of Early Adolescence, 34*(1), 120–151.

Ulbricht, C. M., Chrysanthopoulou, S. A., Levin, L., & Lapane, K. L. (2018). The use of latent class analysis for identifying subtypes of depression: A systematic review. *Psychiatry Research, 266*, 228–246.

United Nations. (2015a). *The Millennium development Goals report 2015.*

United Nations. (2015b). *Transforming our world: The 2030 agenda for sustainable development.* New York: United Nations.

United Nations. (2023). *IAEG-SDGs: Tier classification for global SDG indicators*, 31 March 2023. United Nations Statistical Division. Retrieved 15 August from https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/.

United Nations Statistical Office. (2014). *Guidelines for producing statistics on violence against women: Statistical surveys.* UN.

USAID. Protecting the privacy of DHS survey respondents. https://www.dhsprogram.com/methodology/Protecting-the-Privacy-of-DHS-Survey-Respondents.cfm.

Watson, D., & Parsons, S. (2005). *Domestic abuse of women and men in Ireland: Report on the national study of domestic abuse.*

Weiss, N. H., Dixon-Gordon, K. L., Peasant, C., Jaquier, V., Johnson, C., & Sullivan, T. P. (2017). A latent profile analysis of intimate partner victimization and aggression and examination of between-class differences in psychopathology symptoms and risky behaviors. *Psychological trauma: Theory, Research, Practice, and Policy, 9*(3), 370.

World Health Organization. (2005). *WHO multi-country study on women's health and domestic violence against women: Summary report of initial results on prevalence, health outcomes and women's responses.*

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a monte-carlo study. *Frontiers in Psychology, 5*, 920.

Zhang, Z., Abarda, A., Contractor, A. A., Wang, J., & Dayton, C. M. (2018). Exploring heterogeneity in clinical trials with latent class analysis. *Annals of Translational Medicine, 6*(7).