

BMJ Open Ensuring the comparability of cross-national survey data on intimate partner violence against women: a cross-sectional, population-based study in the European Union

Manuel Martín-Fernández , Enrique Gracia, Marisol Lila

To cite: Martín-Fernández M, Gracia E, Lila M. Ensuring the comparability of cross-national survey data on intimate partner violence against women: a cross-sectional, population-based study in the European Union. *BMJ Open* 2020;**10**:e032231. doi:10.1136/bmjopen-2019-032231

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-032231>).

Received 10 June 2019
Revised 10 January 2020
Accepted 15 January 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Department of Social Psychology, University of Valencia, Valencia, Spain

Correspondence to

Dr Manuel Martín-Fernández; manuel.martin@uv.es

ABSTRACT

Objectives To ensure the cross-national comparability of the set of questions addressing physical and sexual intimate partner violence against women (IPVAW) included in the European Union (EU) Agency for Fundamental Rights (FRA) survey. Once the measurement invariance of these measures is established, we aim to make appropriate and valid comparisons of the levels of physical and sexual IPVAW across the EU countries.

Design Cross-sectional, population-based study.

Participants Data were drawn from the survey conducted by the FRA on violence against women, including the responses of 42 002 adult women from the 28 countries of the EU.

Main outcome measures The set of questions addressing lifetime prevalence of physical and sexual IPVAW used in the FRA survey. The psychometric properties (ie, reliability and validity) of these measures were examined, as well as their latent structure and their measurement invariance across the 28 EU countries.

Results The physical and sexual IPVAW measures presented adequate internal consistency and validity evidence based on their relations to other variables in all countries. A latent two-factor structure was supported and scalar invariance was established across countries. Our results showed that the average levels of physical and sexual IPVAW were highest in Denmark, Finland, Sweden and UK compared with the rest of the EU countries. In many of the other countries the levels of these types of violence overlapped, especially in the case of sexual IPVAW.

Conclusions The findings of this study showed that the set of questions addressing physical and sexual IPVAW included in the FRA survey can be compared across all EU countries, highlighting the importance of testing the measurement equivalence of the instruments used in large sociodemographic surveys in order to make valid cross-national comparisons.

INTRODUCTION

Intimate partner violence against women (IPVAW) has been globally recognised as a major public health problem of epidemic proportions.¹ IPVAW is the form of violence

Strengths and limitations of this study

- This paper is the first to examine the cross-national comparability of survey data addressing physical and sexual intimate partner violence against women (IPVAW) in all European Union (EU) countries.
- Appropriate analyses were applied to examine the psychometric properties, latent structure and the measurement invariance of the measures included in a large sociodemographic survey on IPVAW.
- A latent means analysis was conducted to compare the levels of physical and sexual IPVAW across all EU countries, an approach that takes into account the latent structure of the IPVAW measures and the magnitude of the contribution of each item to the measured construct.
- The cross-sectional design of the survey did not allow for testing measurement invariance across different periods of time.

most commonly suffered by women,²⁻⁴ and it has severe physical and psychological consequences for the well-being of victims and their children, and for society as a whole.⁵⁻⁹

According to the survey conducted by the European Union (EU) Agency for Fundamental Rights (FRA), the estimated prevalence of physical and sexual IPVAW in the EU is 22%, varying across countries from 13% to 32%.¹⁰ One of the main strengths of this survey is that women from the 28 EU countries answered the same set of questions assessing different types of IPVAW. However, the measurement equivalence of these questions across the EU countries has not yet been tested. This hampers the generalisability of these cross-national comparisons, as it not possible to ascertain whether the differences in IPVAW prevalence across all EU member states reflect actual differences between countries, or whether they are the result of

cultural expectations or beliefs about intimate partner violence that may affect the interpretation of the FRA survey questions.

IPVAW can be a culturally sensitive issue, and how it is perceived, conceptualised and interpreted may vary across countries. Thus, to make valid cross-national comparisons based on survey data addressing IPVAW, it is important to test their measurement invariance.^{11–13} Measurement invariance is a crucial prerequisite in cross-national research, since it allows meaningful comparisons across countries ruling out the possibility of cultural bias in the respondents' answers.¹⁴ When measurement invariance is not supported, it means that respondents from different countries interpret and respond differently to the questions, and thus their scores cannot be directly compared.^{15,16} If that is the case, it cannot be assumed that IPVAW is interpreted in the same way in all EU member states, since the same score in one country may reflect a different construct or yield different levels of IPVAW in another. Therefore, computing prevalence rates to compare samples from different countries without first ensuring measurement invariance could lead to unreliable and distorted conclusions, as the validity of such comparisons may be questionable.¹⁷ Previous research has already assessed and established the cross-national comparability of the psychological IPVAW across the EU, but this issue still remains unexplored for physical and sexual forms of this type of violence.¹⁸

The main objective of this study is to ensure the cross-national comparability of the set of questions addressing physical and sexual IPVAW included in the FRA survey, by assessing whether respondents of each country conceptualise and interpret these questions in the same manner. For validity purposes, the relationship of physical and sexual IPVAW to related sociodemographic and background variables (ie, as self-perceived health, household income and experiences of child abuse) was tested.^{19,20} Once the measurement invariance of these measures is established, we aim to make appropriate and valid comparisons of the physical and sexual IPVAW levels across the EU countries.

METHOD

Participants

The data were drawn from the FRA survey on violence against women.¹⁰ This survey includes the responses of 42 002 adult women from the 28 EU countries who were currently or had previously been in an intimate relationship. The responses were collected in person through structured interviews and self-reports, following a two-stage clustered stratified sampling design with equal probability of selection of households within clusters. The average response rate to the survey was 42.1%, ranging from 18.5% in Luxembourg to 84.0% in Hungary. Additional information on sample collection and procedures can be found in the FRA survey technical report.²¹ A special licence for secondary data analysis was requested and granted from FRA (Reference No 102577).

The sample used in this study was selected using the answers from respondents who did not omit any of the survey questions addressing physical and sexual intimate partner violence. The final sample is composed of 39 403 women, aged 18–74 years old, from the 28 EU countries. Sociodemographic information of the sample by country is shown in [table 1](#).

Patient involvement

Participants were not involved in the design, planning and conduct of this study. Participation in this study was anonymous and informed consent was obtained before conducting the interviews.

Measures

The measures used in this study were developed by FRA in collaboration with research experts, survey experts, policymakers, practitioners and non-governmental organisations. All the items included in the questionnaire were translated by experts in the field of IPVAW from English into other languages of the 28 EU member states.²¹

Physical violence

Intimate partner physical violence is addressed in the FRA survey with 10 items describing episodes of physical violence perpetrated by either the current or any previous partner (eg, 'Has your current/previous partner ever slapped you?'). Respondents were asked to indicate how often they had experienced that episode on a 4-point Likert-type scale (1: 'Never', 2: 'Once', 3: '2–5 times', 4: '6 or more times'). The responses to these items were dichotomised (0: 'Never', 1: 'Once or more times'), as in some countries the frequencies of the upper categories were very low (ie, less than 2% when the responses to the '2–5 times' and '6 or more times' categories are collapsed in Austria, Cyprus, Greece, Italy, Latvia, Malta, Slovakia and Slovenia).

Sexual violence

The FRA survey included four items addressing sexual violence committed by the current or any previous partner (eg, 'Has your current/previous partner made you take part in any form of sexual activity when you did not want to or you were unable to refuse?'). The response format of the items was a 4-point Likert-type scale indicating frequency (1: 'Never', 2: 'Once', 3: '2–5 times', 4: '6 or more times'). The responses to these items were also dichotomised (0: 'Never', 1: 'Once or more times').

Validity evidence based on relations to other variables

To test the validity evidence of the physical and sexual IPVAW measures based on relations to other variables,²² the following variables were used:

Self-perceived health

The respondents were asked to respond about their health in general at the beginning of the interview using a 5-point Likert-type graded scale (from 1='Very Bad' to 5='Very Good').

Table 1 Sociodemographic characteristics of the sample in each country (n=39 403)

	n	Age (%)				Self-perceived health		Experiences of child abuse (%)	Income (%)			
		18–24	25–34	35–59	60–74	M	SD		Q1	Q2	Q3	Q4
Austria	1403	7.1	20.0	48.5	24.4	1.77	0.89	30.4	38.5	31.5	21.0	9.0
Belgium	1456	3.5	11.3	52.4	32.8	1.94	0.92	26.7	5.0	28.0	23.1	43.9
Bulgaria	1424	12.5	20.9	44.0	22.7	1.79	0.86	28.0	34.6	23.0	26.4	16.0
Croatia	1438	10.6	16.6	51.2	21.6	2.03	0.87	32.3	11.1	25.0	37.3	26.6
Cyprus	1390	5.2	14.8	53.4	26.6	2.12	0.90	12.4	33.7	28.7	22.9	14.7
Czech Republic	1566	5.2	13.5	44.5	36.8	2.35	1.01	32.6	13.3	37.7	34.4	14.6
Denmark	1453	7.8	17.2	49.9	25.1	2.13	1.09	42.7	26.1	22.4	28.4	23.1
Estonia	1360	18.4	25.9	45.6	10.1	1.65	0.81	48.3	7.9	27.8	45.6	18.7
Finland	1457	10.4	16.6	48.1	24.9	2.07	0.90	52.4	31.3	29.6	22.9	16.2
France	1410	5.7	14.6	57.9	21.8	2.15	0.78	44.4	57.2	32.6	7.3	3.0
Germany	1478	12.0	15.3	48.5	24.1	1.85	0.91	43.0	26.6	33.9	24.9	14.6
Greece	1425	6.9	14.1	46.2	32.9	2.52	0.92	23.3	27.1	22.5	27.4	23.1
Hungary	1451	6.9	15.3	45.2	32.6	2.07	0.81	24.4	21.8	23.3	25.8	29.1
Ireland	1428	6.5	15.8	49.3	28.4	1.92	0.95	28.4	37.9	8.7	9.3	44.2
Italy	1483	7.5	15.9	48.5	28.1	2.63	0.85	34.1	31.8	24.2	21.4	22.5
Latvia	1440	6.0	13.6	55.5	24.8	2.03	0.85	31.9	49.8	32.2	10.8	7.2
Lithuania	1308	6.2	14.1	54.7	25.0	2.13	0.83	17.7	24.2	25.3	29.5	20.9
Luxembourg	889	8.1	11.5	51.1	29.2	2.57	0.85	43.3	37.6	19.7	19.9	22.8
Malta	1365	8.8	16.7	50.8	23.6	1.69	0.89	24.2	21.4	42.3	31.9	4.4
Netherlands	1451	6.2	14.3	57.4	22.0	2.19	0.93	32.9	31.4	25.2	23.0	20.3
Poland	1405	5.2	14.3	50.8	29.7	2.05	0.67	17.8	31.2	39.8	21.2	7.7
Portugal	1424	7.7	15.0	46.4	30.9	2.44	0.98	28.5	20.1	29.2	24.0	26.8
Romania	1465	4.4	12.8	58.4	24.4	2.15	0.77	22.6	19.2	28.8	26.4	25.7
Slovakia	1287	10.3	23.1	46.7	20.0	2.25	1.00	33.9	10.7	24.6	30.3	34.4
Slovenia	1372	5.0	12.5	47.5	34.9	2.61	0.92	13.2	21.0	41.1	22.3	15.6
Spain	1439	9.2	16.6	48.5	25.6	2.42	1.03	29.4	28.7	21.4	20.3	29.6
Sweden	1474	7.5	18.8	52.1	21.6	2.15	0.98	43.8	20.0	23.3	26.1	30.6
UK	1462	6.6	13.7	51.6	28.0	2.13	0.91	37.8	14.7	31.9	33.7	19.6

The percentages and descriptive statistics are unweighted.

Income: Q1=under lowest quartile, Q2=between lowest quartile and median, Q3=between median and highest quartile, Q4=above highest quartile.

M, mean; SD, standard deviation.

Experiences of child abuse

The FRA survey included a set of 11 questions addressing experiences of childhood physical and sexual abuse before the age of 15 (eg, ‘Did an adult who was 18 years or over hit you very hard so that it hurt?’, ‘Did an adult who was 18 years or over expose their genitals to you?’). If any of these questions were answered affirmatively, we considered that the respondent has experienced child abuse during their childhood.

Income

Reported income was measured in each country as household income quartiles (ie, ‘below lowest quartile’,

‘between lowest quartile and median’, ‘between median and highest quartile’, ‘above highest quartile’).

Statistical analyses

The main purpose of this study was to ensure the measurement equivalence of the physical and sexual violence measures of the FRA survey across the 28 EU member states in order to make appropriate and valid comparisons between these countries. To this end, the psychometric properties (ie, reliability and validity) and the latent structure of these measures was assessed, and then measurement invariance was tested across the 28 EU countries.

First, the mean, SD, skew and kurtosis statistics were obtained. The internal consistency of the physical and sexual violence measures in each country was evaluated by computing the Cronbach's α and McDonald's ω , and using the correlation of each item with the rest of its set of questions (ie, item–test corrected correlation).

Second, a confirmatory factor analysis (CFA) was conducted to determine the latent structure of the two IPVAV measures in each country. Two models were compared: a one-factor model, in which all the items load onto a single intimate partner violence factor, and a two-factor model, setting one factor for the physical violence items and another one for the sexual violence items. Given the categorical nature of the items, weighted least squares with adjusted means and variances (WLSMV) was used as estimation method.²³ Model fit was evaluated using a combination of fit indices: comparative fit index (CFI), Tucker-Lewis index (TLI) and root mean square error of approximation (RMSEA). CFI and TLI values ≥ 0.95 , and RMSEA values ≤ 0.06 are indicative of good fit.^{24 25}

After establishing the factor model, a series of multi-group CFA (MG-CFA) were carried out, testing three different levels of measurement invariance: configural, metric and scalar.^{12 26} Configural invariance applies the same factor model to all groups (ie, countries), with no equality constraints for any parameters, ensuring that the same conceptualisation of the construct is supported across countries. Metric invariance specifies that the values of the factor loadings are equal across countries, implying that each item is contributing equally and having a similar degree of importance on the factor in all the countries. Scalar invariance holds both factor loadings and item thresholds to be invariant, ensuring that respondents from different countries with the same response pattern on the items will yield the same factor score. If scalar invariance is supported, then the scores on the physical and sexual IPVAV measures can be compared across countries.

To assess which of these invariance levels was best supported by the data, the change in the CFI (Δ CFI) and RMSEA (Δ RMSEA) fit indices was computed, following the general guidelines of Cheung and Rensvold.²⁷ These guidelines, however, were developed for continuous data, and as Sass *et al* noted,²⁸ the interpretation of these fit indices should be taken with caution in the case of categorical data. For this reason, instead of Chen's usual cut-off values for the changes in CFI and RMSEA (ie, Δ CFI ≤ 0.010 and Δ RMSEA ≤ 0.015),²⁹ we used the cut-off values proposed by Meade *et al*: Δ CFI ≤ 0.002 and Δ RMSEA ≥ -0.007 .³⁰ These cut-offs are more restrictive and tend to perform as well as maximum likelihood-based procedures when the sample size is large and the items are not normally distributed.²⁸ This approach has also been used to assess the measurement invariance of the set of questions addressing psychological IPVAV in the FRA survey.¹⁸

Having established an invariant factor model, validity evidence based on relations to other variables was tested using the invariant factor scores of the IPVAV measures.

A one-way analysis of variance was computed for both physical and sexual violence measures of the FRA survey, testing differences by self-perceived health, experiences of child abuse and income. These analyses were conducted for the total sample to avoid type I errors. The size effect of the variables was assessed with the partial eta-squared statistic, using values above 0.01, 0.06 and 0.14, as indicative of small, medium and large size effects, respectively.³¹

Finally, after determining that both physical and sexual IPVAV measures were psychometrically sound and share an invariant latent structure in all EU countries, the factor means of physical and sexual IPVAV were compared across countries conducting a latent means analysis. This procedure is more appropriate and statistically sophisticated than simply computing the average prevalence for each country, as it does not assume that all the items have the same relevance to assess the construct and, moreover, it takes into account all the constraints for the invariant measurement model.^{16 17 26} The magnitude of these cross-national comparisons was evaluated using Cohen's *d* statistic, indicating *d* values above 0.20, 0.50 and 0.80 for small, medium and large size effects, respectively.³² This statistic can be used to obtain the Cohen's U_3 statistic, which indicates the percentage of cases of one country that is higher than the average of another.^{33–35} The Cohen's *d* and U_3 statistics of the comparisons between each pair of countries can be found in the online supplementary material 1.

Descriptive analyses, internal consistency and validity analyses were conducted using the statistical package R and the *psych* library.^{36 37} CFA and MG-CFA were computed using *Mplus 7*.³⁸

RESULTS

Descriptive analysis and internal consistency

Descriptive statistics of the physical and sexual violence measures are presented in table 2. All the items showed mean values close to zero, and high skew and kurtosis values, especially in the sexual violence measure, indicating that most of the respondents reported having no experience of the episodes described by the items. The corrected item–test correlations were in general high, pointing to a strong relationship between the items and the rest of the questions. Regarding the internal consistency, both measures showed an adequate internal consistency in the complete data set, with α and ω values above 0.70. Separating the sample by countries yields similar results in both the physical violence and sexual violence measures (ranging, respectively, from $\alpha=0.84$ in Sweden to $\alpha=0.90$ in Ireland and Lithuania, and from $\alpha=0.69$ in Slovenia to $\alpha=0.91$ in the Netherlands).

The items 7 and 9 of the physical violence measure presented an extremely low variance (ie, 'being burned', and 'being cut, stabbed or shot'), below 0.01. Given this lack of variability, we decided to remove these items from this measure. As a result, the Cronbach's α of the physical violence measure increased to 0.89 (ranging from $\alpha=0.84$

Table 2 Psychometric properties of the physical violence and sexual violence measures

	M	SD	Skew	Kurtosis	r _{item-test}
Physical violence (Cronbach's $\alpha=0.87$, McDonald's $\omega=0.90$)					
Threatened to hurt you physically	0.14	0.35	2.05	2.18	0.70
Pushed you or shoved you	0.16	0.36	1.88	1.53	0.73
Slapped you	0.13	0.33	2.23	2.99	0.73
Thrown a hard object at you	0.06	0.24	3.66	11.43	0.65
Grabbed you or pulled your hair	0.08	0.28	3.03	7.18	0.74
Beat you with a fist or a hard object, or kicked you	0.07	0.25	3.50	10.26	0.72
Burned you	0.01	0.09	11.55	131.34	0.27
Tried to suffocate you or strangle you	0.03	0.18	5.25	25.60	0.53
Cut or stabbed you, or shot at you	0.01	0.09	11.37	127.23	0.29
Beat your head against something	0.04	0.18	5.03	23.32	0.60
Sexual violence (Cronbach's $\alpha=0.85$, McDonald's $\omega=0.88$)					
Forced you into sexual intercourse by holding you down or hurting you in some way	0.04	0.19	4.93	22.34	0.73
Attempted to force you into sexual intercourse by holding you down or hurting you in some way	0.04	0.19	4.94	22.44	0.73
Made you take part in any form of sexual activity when you did not want to or you were unable to refuse	0.04	0.19	4.94	22.42	0.66
Consented to sexual activity because you were afraid of what your current partner might do if you refused	0.05	0.21	4.28	16.35	0.68

Skew and kurtosis SE was below 0.01.

M, mean; r_{item-test}, corrected item–test correlation; SD, standard deviation.

in Slovenia to $\alpha=0.92$ in Romania and Lithuania), and McDonald's ω increased to 0.92.

Confirmatory factor analysis

Two CFA solutions were tested separately in each country: a one-factor model, assuming that all the items load onto a single IPVAV factor, and a two-factor model, setting one separate factor for the physical and sexual violence items (table 3). The CFI and TLI indices were excellent for both the one and two-factor models, with values around 0.98 and 0.99 in all 28 EU countries, respectively. However, the residuals for the one-factor model were above the RMSEA ≤ 0.06 cut-off criterion for a well-fitting model in Belgium, Denmark, Finland, France, Germany, Latvia, the Netherlands, Portugal, Sweden and UK (the highest was Denmark, with 0.077, and the closest to the cut-off was Portugal, with 0.061), whereas in the two-factor model the RMSEA was below this cut-off (the highest was 0.031 in Denmark and the lowest, 0.011 in Poland). For this reason, we kept the two-factor model as the latent structure of the physical and sexual IPVAV measures.

Measurement invariance

The analysis of measurement invariance supported the configural, metric and scalar invariance levels for the physical and sexual violence measures across all EU countries (table 4). Attending to the fit indices, when the loadings were fixed to have the same value across all countries they did not differ substantially from the configural model

($\Delta\text{CFI}=0.002$, $\Delta\text{RMSEA}=-0.006$), supporting the metric invariance level. In the same direction, constraining the item thresholds as well as the item loadings did not substantially reduce the fit of the model ($\Delta\text{CFI}=0.002$, $\Delta\text{RMSEA}=-0.006$), indicating that the scalar invariance level could hold.

The resulting item parameters are displayed in figure 1. All the items presented high standardised loadings, indicating a strong relationship of the items to the factor. The correlation between factors varies across countries, as the factor covariance matrix was freed in each country. The value for this correlation was around 0.75 in most of the countries, ranging from 0.56 in Denmark to 0.91 in Croatia. The model parameters for all countries can be found in the online supplementary material 2.

Validity evidence based on relations to other variables

The factor scores of the invariant model were used for the validity analyses in all countries. Significant and substantive differences were found in the physical IPVAV scores when self-perceived health ($F(4)=94.0$, $p<0.001$, $\eta^2=0.01$), experiences of child abuse ($F(1)=984.7$, $p<0.001$, $\eta^2=0.03$) and income ($F(1)=127.3$, $p<0.001$, $\eta^2=0.01$) were taken into account, showing that women with higher scores on the physical IPVAV factor had lower self-perceived health levels, lower income and more experiences of child abuse.

In the case of the sexual IPVAV factor scores, substantive differences were also found in self-perceived health

Table 3 Confirmatory factor analysis fit indices by country

	One-factor model			Two-factor model		
	CFI	TLI	RMSEA	CFI	TLI	RMSEA
Austria	0.991	0.989	0.047	0.999	0.998	0.019
Belgium	0.986	0.982	0.063	0.999	0.999	0.018
Bulgaria	0.994	0.993	0.058	0.999	0.998	0.026
Croatia	0.995	0.993	0.038	0.999	0.999	0.017
Cyprus	0.997	0.996	0.044	0.999	0.999	0.014
Czech Republic	0.983	0.979	0.060	0.997	0.996	0.026
Denmark	0.965	0.957	0.077	0.994	0.993	0.031
Estonia	0.990	0.988	0.056	0.999	0.999	0.017
Finland	0.981	0.977	0.061	0.999	0.999	0.004
France	0.981	0.977	0.064	0.998	0.998	0.021
Germany	0.980	0.975	0.062	0.998	0.998	0.019
Greece	0.993	0.992	0.051	0.998	0.998	0.027
Hungary	0.991	0.989	0.052	0.998	0.997	0.025
Ireland	0.997	0.996	0.050	0.999	0.999	0.020
Italy	0.990	0.988	0.046	0.999	0.998	0.017
Latvia	0.981	0.977	0.073	0.997	0.996	0.029
Lithuania	0.996	0.995	0.060	0.999	0.999	0.019
Luxembourg	0.991	0.988	0.049	0.999	0.999	0.005
Malta	0.989	0.987	0.045	0.999	0.999	0.007
Netherlands	0.986	0.983	0.071	0.999	0.999	0.015
Poland	0.996	0.995	0.045	0.999	0.999	0.011
Portugal	0.992	0.991	0.061	0.999	0.999	0.020
Romania	0.997	0.996	0.056	0.999	0.999	0.026
Slovakia	0.991	0.989	0.055	0.998	0.998	0.023
Slovenia	0.996	0.995	0.026	0.999	0.999	0.008
Spain	0.992	0.990	0.043	0.998	0.997	0.023
Sweden	0.967	0.959	0.074	0.995	0.994	0.029
UK	0.991	0.989	0.062	0.999	0.999	0.020

CFI, comparative fit index; RMSEA, root mean square error of approximation; TLI, Tucker-Lewis index.

($F(4)=82.6$, $p<0.001$, $\eta^2=0.01$), experiences of child abuse ($F(1)=311.4$, $p<0.001$, $\eta^2=0.03$) and income ($F(1)=113.4$, $p<0.001$, $\eta^2=0.01$). These differences were in the same direction as in the physical IPVAV, as the women with higher scores on the sexual IPVAV factor presented lower self-perceived health levels, lower income and more experiences of child abuse.

Latent means analysis

After determining that the physical and sexual IPVAV measures are psychometrically sound, and after ensuring their measurement equivalence across all EU countries, the means of the physical and sexual violence factors can now be compared by conducting an MG-CFA. Denmark was used as the reference group, as it was one of the

Table 4 Measurement invariance fit indices

Model	χ^2	df	CFI	TLI	RMSEA (95% CI)
Configural	2461.08	1484	0.999	0.998	0.022 (0.020 to 0.023)
Metric	3819.28	1808	0.997	0.997	0.028 (0.027 to 0.029)
Scalar	5530.65	2078	0.995	0.996	0.034 (0.033 to 0.035)

χ^2 , adjusted χ^2 test for model fit; CFI, comparative fit index; df, degrees of freedom; RMSEA, root mean square error of approximation; TLI, Tucker-Lewis index.

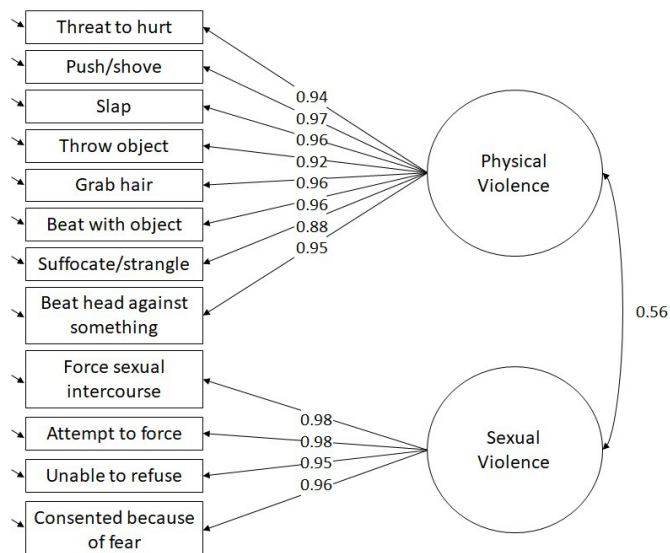


Figure 1 Multigroup confirmatory factor analysis: scalar invariance model. The standardised factor loadings belong to the reference country (ie, Denmark). As in the standardised solution the factor variances are fixed to 1, there are slight differences in the decimals of the loadings in each country.

countries presenting the highest levels of physical and sexual IPVAW. The factor means for Denmark were fixed to zero and their variances fixed to one. In the rest of the countries these parameters were freely estimated. In this way, the estimated mean parameter of the standardised solution represents the difference in factor means between each country and Denmark. The resulting model presented a very good fit to the data (CFI=0.997, TLI=0.997, RMSEA=0.027 (0.026–0.029)). The standardised factor means for each country are displayed in figures 2 and 3.

Although Denmark was the country with the highest levels of physical IPVAW, the differences between this country and Latvia ($z=-0.01, p=0.850, d=0.01$), Finland ($z=-0.04, p=0.587, d=0.04$), UK ($z=-0.12, p=0.059, d=0.13$) and Sweden ($z=-0.12, p=0.115, d=0.12$) were negligible, with Cohen’s d values below 0.20. Large substantive differences were found between Denmark and Ireland ($z=-0.70, p<0.001, d=0.85$), Slovenia ($z=-0.75, p<0.001, d=0.84$), Spain ($z=-0.75, p<0.001, d=0.86$), Cyprus ($z=-0.77, p<0.001, d=0.94$) and Poland ($z=-0.79, p<0.001, d=0.96$).

It should be noted that the CIs of the physical IPVAW factor means overlapped between many of the EU countries, indicating no significant differences between them. However, the countries could be grouped according to the CIs. Thus, countries whose factor mean falls within the CI of Latvia showed on average higher levels of physical IPVAW (Denmark, Finland, UK and Sweden). On the other hand, countries with factor means falling within the CI of Poland—the country with the lowest factor mean—presented lower levels of physical IPVAW (Ireland, Slovenia, Spain and Cyprus). The rest of the countries fell somewhere in between, with some countries like Lithuania or the Netherlands closer to the countries with higher levels of physical IPVAW, and others such as Greece or Italy closer to the countries with lower levels.

Regarding the sexual IPVAW measure, the countries with higher latent means in this factor were, along with Denmark, Finland ($z=0.00, p=0.986, d=0.00$), Sweden ($z=-0.10, p=0.537, d=0.10$), Luxembourg ($z=-0.31, p=0.082, d=0.34$), Bulgaria ($z=-0.32, p=0.072, d=0.34$) and Slovenia ($z=-0.34, p=0.130, d=0.32$). Large differences in the sexual IPVAW factor were found between Denmark and Lithuania ($z=-0.87, p<0.001, d=1.06$), Croatia ($z=-0.88, p<0.001, d=1.04$), Austria ($z=-0.90, p<0.001,$

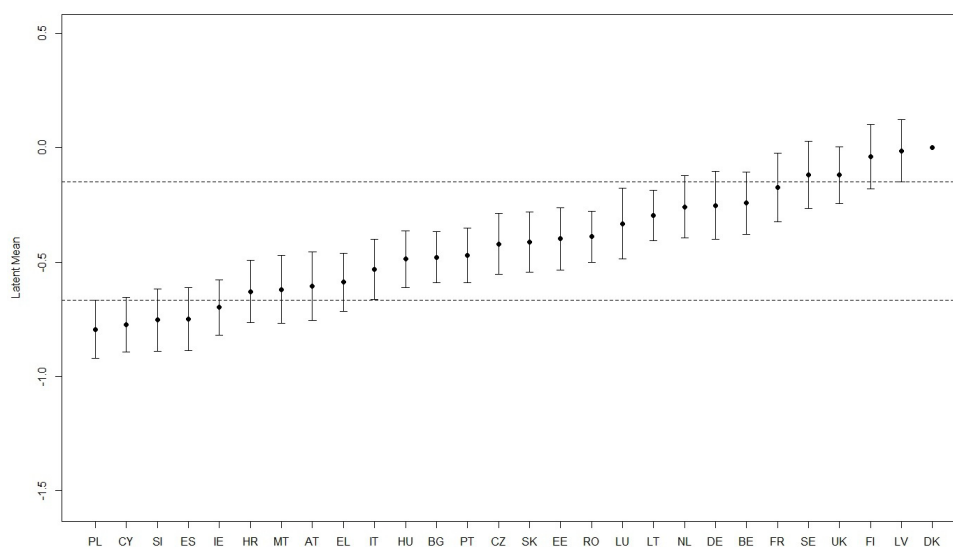


Figure 2 Physical violence latent means across the EU. AT, Austria; BE, Belgium; BG, Bulgaria; CY, Cyprus; CZ, Czech Republic; DE, Germany; DK, Denmark; EE, Estonia; EL, Greece; ES, Spain; FI, Finland; FR, France; HR, Croatia; HU, Hungary; IE, Ireland; IT, Italy; LT, Lithuania; LU, Luxembourg; LV, Latvia; MT, Malta; NL, Netherlands; PL, Poland; PT, Portugal; RO, Romania; SE, Sweden; SI, Slovenia; SK, Slovakia; UK, United Kingdom. Denmark is the reference country.

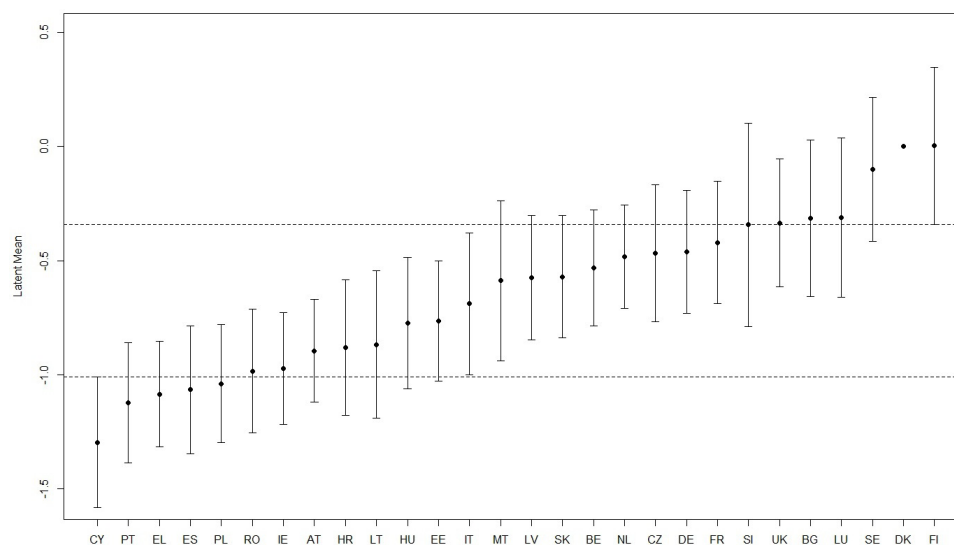


Figure 3 Sexual violence factor means across the EU. AT, Austria; BE, Belgium; BG, Bulgaria; CY, Cyprus; CZ, Czech Republic; DE, Germany; DK, Denmark; EE, Estonia; EL, Greece; ES, Spain; FI, Finland; FR, France; HR, Croatia; HU, Hungary; IE, Ireland; IT, Italy; LT, Lithuania; LU, Luxembourg; LV, Latvia; MT, Malta; NL, Netherlands; PL, Poland; PT, Portugal; RO, Romania; SE, Sweden; SI, Slovenia; SK, Slovakia; UK, United Kingdom. Denmark is the reference country.

$d=1.13$), Ireland ($z=-0.97$, $p<0.001$, $d=1.29$), Romania ($z=-0.98$, $p<0.001$, $d=1.24$), Poland ($z=-1.04$, $p<0.001$, $d=1.31$), Spain ($z=-1.07$, $p<0.001$, $d=1.32$), Greece ($z=-1.08$, $p<0.001$, $d=1.39$), Portugal ($z=-1.12$, $p<0.001$, $d=1.43$) and Cyprus ($z=-1.97$, $p<0.001$, $d=1.71$).

The CIs of the sexual IPVAV factor means were, however, much wider than in the case of physical IPVAV, as this measure comprised only four items. As a result, most of the countries' factor means CI greatly overlapped. In this case, the CI of Finland was used to group the countries that tend to present higher levels of sexual IPVAV (Denmark, Sweden, Belgium, Luxembourg, UK and Slovenia), whereas the CI of Cyprus was used to delimit those countries with lower levels in this factor (Portugal, Greece, Spain and Poland). The rest of the countries fall into the intermediate area, with some countries such as France and Germany nearer to the countries with higher levels of sexual IPVAV, and others like Romania and Ireland closer to the countries with lower levels of this type of violence. Italy was in the middle of the distribution, and was the only country whose CI did not overlap with the intervals for either Finland or Cyprus.

Comparisons of the physical and sexual IPVAV factor scores can also be made between each pair of countries. For instance, we compare Finland—one of the countries with the highest levels of physical and sexual IPVAV—with the Netherlands and Spain—countries with moderate and low levels of both types of IPVAV, respectively. The differences between Finland and the Netherlands were small in the case of physical IPVAV, showing that 59.9% of the Finnish sample had higher values in this factor than the average of the Dutch sample ($d=0.24$, $U_3=0.595$). Medium differences were found in the case of sexual IPVAV, as 72.6% of the Finnish sample presented higher scores in the sexual IPVAV factor than the average

of the Dutch sample ($d=0.60$, $U_3=0.726$). When Finland is compared with Spain, large differences were found in both factors: 79.1% of the Finnish sample presented higher values in the physical IPVAV factor than the average of the Spanish sample ($d=0.81$, $U_3=0.791$), and 90.7% in the case of sexual IPVAV ($d=1.32$, $U_3=0.907$). The comparisons between each pair of countries together with their associated size effect can be found in the online supplementary material 1.

DISCUSSION

In this measurement invariance study, we conducted a set of analyses to ensure the cross-national comparability of the questions addressing physical and sexual IPVAV used in the FRA survey, and examined how physical and sexual IPVAV levels were distributed across all EU countries.

The first set of analyses aimed to study the psychometric properties (ie, reliability and validity) of the questions addressing physical and sexual IPVAV across the EU. Our results showed that the measures of the FRA survey were measuring two different constructs: physical and sexual IPVAV. The two factors were related and showed an adequate internal consistency in all EU countries (with α and ω values above 0.70 in almost all countries). Regarding the validity evidence based on relations with other variables, we found, as expected, that women who reported having experiences of child abuse were more likely to show higher levels of physical and sexual IPVAV. This finding is consistent with previous research showing that women who were victimised in childhood have a higher risk of being victimised during adulthood.^{39–44} In addition, we found that the women presenting higher levels of physical and sexual IPVAV also reported

lower levels of self-perceived health and income, a well-established finding in the IPVAV literature.^{45–49}

The second set of analyses aimed to address the comparability of the physical and sexual IPVAV measures across the 28 EU countries, to rule out the possibility of measurement bias. To do so, a series of MG-CFA were conducted to test and establish measurement invariance. Although previous studies have acknowledged the difficulties in establishing the metric or scalar invariance levels when several groups are used,^{14 50 51} we were able to demonstrate scalar invariance (ie, equal factor loadings and thresholds) across the 28 EU countries. This means that it is possible to exclude the country as a biasing variable, which in turn suggests that the respondents conceptualised and interpreted the physical and sexual IPVAV measures included in the FRA survey in a similar way across countries.^{52 53} This finding may be due to the type of questions included in the FRA survey, as they are mostly behaviourally oriented (eg, being stabbed, cut, slapped, or being forced into sexual intercourse).

Once an invariant model had been established for the physical and sexual IPVAV measures, we were able to make proper comparisons between countries. Instead of computing the prevalence using the set of questions from the FRA survey, an approach that ignores the latent structure of the IPVAV measures and the magnitude of the contribution of each item to the construct, we conducted a latent means analysis, comparing the factor means of each country.^{12 15 16} This is one of the main strengths of this study, and our findings suggest that there were almost no differences between many of the EU countries, since their CIs overlapped considerably, implying that the levels of sexual and physical IPVAV were fairly similar between them. There were, however, substantial differences between the countries with higher levels of physical and sexual IPVAV and those with lower levels. In particular, countries such as Denmark, Finland, Sweden and UK showed on average higher levels than countries like Cyprus, Poland and Spain. These findings are in line with what is known as the Nordic paradox, as Denmark, Finland and Sweden—countries with the highest levels of gender equality in the EU—were among the countries with highest levels of physical and sexual IPVAV.^{54 55}

This study has some limitations. First, we used WLSMV as the estimation method for the measurement invariance analysis due to the asymmetrical distribution of the responses to the set of questions included in the FRA survey. According to Sass *et al*,²⁸ the use of fit indices with this method to test measurement invariance could lead to higher rates of type I errors, assuming an invariant model when actually the instrument is non-invariant. To address this issue, we decided to use the cut-offs proposed by Meade *et al* for the CFI and RMSEA,³⁰ as they are more restrictive than those of Chen.²⁹ Second, in this case, χ^2 -based tests to compare the different levels of measurement invariance could not be used in conjunction with the fit indices, since this statistic is known to be sensitive to large sample sizes. Third, the cross-sectional design of the

survey did not allow for testing measurement invariance across different periods of time. Fourth, the wide range of the CIs of the physical and sexual IPVAV factor means suggests that the measures included in the FRA survey could be improved in order to obtain more accurate and reliable estimations of the IPVAV levels, especially in the case of sexual IPVAV. Fifth, self-selection bias is another potential issue, since only a 42% of the respondents agreed initially to be interviewed. This is also reflected in the wide variability in the response rates across the EU countries, with some countries presenting response rates lower than 30% (eg, Luxembourg, Netherlands or Sweden), whereas others showed response rates above 60% (eg, Cyprus, Hungary or Latvia).²⁴ Finally, future research could also examine the relationship of physical and sexual IPVAV to other IPVAV measures included in the survey, such as psychological IPVAV, sexual harassment or stalking behaviour.¹⁸

Taken together, the results of this study showed that the set of questions addressing physical and sexual IPVAV included in the FRA survey can be compared across all EU countries. Although more research is still needed to assess the relevance that the country may have in accounting for the differences of physical and sexual IPVAV across the EU,⁵⁶ this study constitutes an important advance towards a more accurate evaluation of cross-national differences in physical and sexual IPVAV, highlighting the importance of testing the measurement equivalence of the instruments used in large sociodemographic surveys in order to make valid cross-national comparisons.

Contributors MMF designed the analytic strategy, conducted the statistical analysis and wrote the manuscript. EG conceived the study, supervised the writing of the manuscript and acquired the funding and the FRA special licence. ML supervised the writing of the manuscript and acquired the funding.

Funding This research was supported by grants PSI2017-84764-P (EG, ML) and BES-2015-075576 (MMF) from the Spanish Ministry of Economy and Competitiveness (<http://www.mineco.gob.es/>).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The European Union Agency for Fundamental Rights supervised and approved the secondary data analyses conducted in this study, as well as the study design, granting a special licence for this purpose (Reference No 102577).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data sets used for the results presented in this manuscript are available from the UK Data Service (<https://www.ukdataservice.ac.uk>). The data set details are the following: Title: European Union Agency for Fundamental Rights: Violence Against Women Survey, 2012: Special Licence Access. Alternative title: FRA VAW Survey. Persistent identifier: 10.5255/UKDA-SN-7730-1.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Manuel Martín-Fernández <http://orcid.org/0000-0003-3606-3559>

REFERENCES

- 1 World Health Organization. *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. Geneva: World Health Organization, 2013.
- 2 Devries KM, Mak JYT, García-Moreno C, et al. Global health: the global prevalence of intimate partner violence against women. *Science* 2013;340:1527–8.
- 3 Garcia-Moreno C, Jansen HAFM, Ellsberg M, et al. Prevalence of intimate partner violence: findings from the who multi-country study on women's health and domestic violence. *The Lancet* 2006;368:1260–9.
- 4 Stöckl H, Devries K, Rotstein A, et al. The global prevalence of intimate partner homicide: a systematic review. *The Lancet* 2013;382:859–65.
- 5 Campbell JC. Health consequences of intimate partner violence. *The Lancet* 2002;359:1331–6.
- 6 Craparo G, Gori A, Petruccioli I, et al. Intimate partner violence: relationships between alexithymia, depression, attachment styles, and coping strategies of battered women. *J Sex Med* 2014;11:1484–94.
- 7 Ellsberg M, Jansen HAFM, Heise L, et al. Intimate partner violence and women's physical and mental health in the who multi-country study on women's health and domestic violence: an observational study. *The Lancet* 2008;371:1165–72.
- 8 Guedes A, Bott S, Garcia-Moreno C, et al. Bridging the gaps: a global review of intersections of violence against women and violence against children. *Glob Health Action* 2016;9:e31516.
- 9 Vilarinho M, Amado BG, Vázquez MJ, et al. Psychological harm in women victims of intimate partner violence: epidemiology and quantification of injury in mental health markers. *Psychosoc Interv* 2018;27:145–52.
- 10 European Union Agency for Fundamental Rights. *Violence against women: an EU-wide survey. main results*. Luxembourg: Publications Office of the European Union, 2014.
- 11 Gregorich SE. Do self-report instruments allow meaningful comparisons across diverse population groups? testing measurement invariance using the confirmatory factor analysis framework. *Med Care* 2006;44:s78–94.
- 12 L. Milfont T, Fischer R. Testing measurement invariance across groups: applications in cross-cultural research. *Int J Psychol Res* 2010;3:111–30.
- 13 Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Methods* 2000;3:4–70.
- 14 Davidov E, Meuleman B, Cieciuch J, et al. Measurement equivalence in cross-national research. *Annu Rev Sociol* 2014;40:55–75.
- 15 Jang S, Kim ES, Cao C, et al. Measurement invariance of the satisfaction with life scale across 26 countries. *J Cross Cult Psychol* 2017;48:560–76.
- 16 Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev Rev* 2016;41:71–90.
- 17 van de Schoot R, Lugtig P, Hox J. A checklist for testing measurement invariance. *Eur J Dev Psychol* 2012;9:486–92.
- 18 Martín-Fernández M, Gracia E, Lila M. Psychological intimate partner violence against women in the European Union: a cross-national invariance study. *BMC Public Health* 2019;19:1739.
- 19 Staggs SL, Riger S. Effects of intimate partner violence on low-income women's health and employment. *Am J Community Psychol* 2005;36:133–45.
- 20 Thompson MP, Arias I, Basile KC, et al. The association between childhood physical and sexual victimization and health problems in adulthood in a nationally representative sample of women. *J Interpers Violence* 2002;17:1115–29.
- 21 European Union Agency for Fundamental Rights. *Violence against women: an EU-wide survey: survey methodology, sample and fieldwork*. Luxembourg: Publications Office of the European Union, 2014.
- 22 American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association, 1999.
- 23 Beauducél A, Herzberg PY. On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal* 2006;13:186–203.
- 24 LT H, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:1–55.
- 25 MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods* 1996;1:130–49.
- 26 Sass DA. Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *J Psychoeduc Assess* 2011;29:347–63.
- 27 Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal* 2002;9:233–55.
- 28 Sass DA, Schmitt TA, Marsh HW. Evaluating model fit with ordered categorical data within a measurement invariance framework: a comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal* 2014;21:167–80.
- 29 Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal* 2007;14:464–504.
- 30 Meade AW, Johnson EC, Braddy PW. Power and sensitivity of alternative fit indices in tests of measurement invariance. *J Appl Psychol* 2008;93:568–92.
- 31 Miles J, Shevlin M. *Applying regression and correlation: a guide for students and researchers*. London: Sage, 2001.
- 32 Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates, 1988.
- 33 Grissom RJ, Kim JJ. *Effect sizes for research: univariate and multivariate applications*. 2nd ed. New York: Routledge, 2012.
- 34 Ruscio J. A probability-based measure of effect size: robustness to base rates and other factors. *Psychol Methods* 2008;13:19–30.
- 35 Hanel PHP, Mehler DMA. Beyond reporting statistical significance: identifying informative effect sizes to improve scientific communication. *Public Underst Sci* 2019;28:468–85.
- 36 R Core Team. R: a language and environment for statistical computing. R foundation for statistical computing, 2019. Available: <https://www.R-project.org/>
- 37 Revelle W. *Psych: procedures for personality and psychological research*. Evanston, Illinois, USA: Northwestern University, 2018. <http://CRAN.R-project.org/package=psychVersion=1.8.12>
- 38 Muthén LK, Muthén BO. *Mplus user's guide*. Los Angeles, CA: Author, 2010.
- 39 Appel AE, Holden GW. The co-occurrence of spouse and physical child abuse: a review and appraisal. *Journal of Family Psychology* 1998;12:578–99.
- 40 Capaldi DM, Knoble NB, Shortt JW, et al. A systematic review of risk factors for intimate partner violence. *Partner Abuse* 2012;3:231–80.
- 41 Classen CC, Palesh OG, Aggarwal R. Sexual revictimization: a review of the empirical literature. *Trauma Violence Abuse* 2005;6:103–29.
- 42 Margolin G, Gordis EB, Medina AM, et al. The co-occurrence of husband-to-wife aggression, family-of-origin aggression, and child abuse potential in a community sample. *J Interpers Violence* 2003;18:413–40.
- 43 Schumacher JA, Feldbau-Kohn S, Smith Slep AM, et al. Risk factors for male-to-female partner physical abuse. *Aggress Violent Behav* 2001;6:281–352.
- 44 Till-Tentschert U. The Relation Between Violence Experienced in Childhood and Women's Exposure to Violence in Later Life: Evidence From Europe. *J Interpers Violence* 2017;32:1874–94.
- 45 Coker AL, Davis KE, Arias I, et al. Physical and mental health effects of intimate partner violence for men and women. *Am J Prev Med* 2002;23:260–8.
- 46 Lövestad S, Krantz G. Men's and women's exposure and perpetration of partner violence: an epidemiological study from Sweden. *BMC Public Health* 2012;12:e945.
- 47 Nybergh L, Taft C, Krantz G. Psychometric properties of the who violence against women instrument in a female population-based sample in Sweden: a cross-sectional survey. *BMJ Open* 2013;3:e002053.
- 48 Reichel D. Determinants of intimate partner violence in Europe: the role of socioeconomic status, inequality, and partner behavior. *J Interpers Violence* 2017;32:1853–73.
- 49 Sokoloff NJ, Dupont I. Domestic violence at the intersections of race, class, and gender: challenges and contributions to understanding violence against marginalized women in diverse communities. *Violence Against Women* 2005;11:38–64.
- 50 Davidov E, Schmidt P, Schwartz SH. Bringing values back in: the adequacy of the European social survey to measure values in 20 countries. *Public Opin Q* 2008;72:420–45.
- 51 Muthén B, Asparouhov T. Recent methods for the study of measurement invariance with many groups: alignment and random effects. *Social Method Res* 2018;47:637–64.
- 52 Bialosiewicz S, Murphy K, Berry T. *An introduction to measurement invariance testing: resource packet for participants*. 2018. Claremont, USA: Claremont Evaluation Center, 2013. <http://comm>.

- eval.org/HigherLogic/System/DownloadDocumentFile.ashx? DocumentFileKey=63758fed-a490-43f2-8862-2de0217a08b8
- 53 Lubke GH, Dolan CV, Kelderman H, *et al.* Weak measurement invariance with respect to unmeasured variables: an implication of strict factorial invariance. *Br J Math Stat Psychol* 2003;56:231–48.
- 54 Gracia E, Merlo J. Intimate partner violence against women and the Nordic paradox. *Soc Sci Med* 2016;157:27–30.
- 55 Gracia E, Martín-Fernández M, Lila M, *et al.* Prevalence of intimate partner violence against women in Sweden and Spain: A psychometric study of the ‘Nordic paradox’. *PLoS One* 2019;14:e0217015.
- 56 Ivert A-K, Gracia E, Lila M, *et al.* Does country-level gender equality explain individual risk of intimate partner violence against women? A multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA) in the European Union. *Eur J Public Health* 2019;385.