

Bettio, Francesca; Ticci, Elisa; Betti, Gianni

Working Paper

A fuzzy approach to measuring violence against women and its severity

GLO Discussion Paper, No. 171

Provided in Cooperation with:
Global Labor Organization (GLO)

Suggested Citation: Bettio, Francesca; Ticci, Elisa; Betti, Gianni (2018) : A fuzzy approach to measuring violence against women and its severity, GLO Discussion Paper, No. 171, Global Labor Organization (GLO), Maastricht

This Version is available at:
<http://hdl.handle.net/10419/173222>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A fuzzy approach to measuring violence against women and its severity

Francesca Bettio[§], Elisa Ticci⁺ and Gianni Betti[§]

Abstract

We develop a scale of severity of violence against women based on fuzzy set theory. The scale can be used to derive fuzzy indexes of violence which account for the prevalence, frequency and severity of violence. Using the results of the survey conducted by the European Agency for Human Rights (FRA) we find strong congruence of ranking between the proposed scale and three widely used alternatives – the Conflict Tactic Scale, The Severity of Violence Against Women Scale and the Index of Spouse Abuse. Unlike existing alternatives, however, the scale that we propose is based on objective information rather than subjective assessment; it is parsimonious in terms of the amount of information that it requires; and it is less vulnerable to risks of cultural bias. As an example of the uses to which fuzzy measurement of violence can be put, we compute fuzzy indexes of intimate partner violence for European countries and find a clear, inverse correlation across countries with the degree of gender equality.

Keywords

violence against women, fuzzy set theory, severity of violence scales, Europe, gender

JEL codes

C49, D63, J16

(draft January 2018)

[§]*University of Siena, Department of Economics and Statistics*

⁺*University of Siena, Department of Political and International Sciences*

1. Overview and motivation

Violence against women (WAW) continues to attract considerable attention in social and medical sciences, including disciplines like economics where issues of violence have never been part of core disciplinary interests. Increasingly sophisticated and large scale surveys have spread awareness that such violence is common enough not to be dismissed as an occasional accident or something experienced by a minority of the population. For example, according to the most extensive European survey carried out to date on violence against women in Europe, some 31% of women experienced one or more acts of physical violence since the age of 15 (Fundamental Rights Agency, 2014: 21; FRA survey henceforth). Our motivation for this article stemmed, in fact, from analysis of the results of FRA survey (Bettio and Ticci, 2017). The survey administered a comprehensive face-to-face questionnaire to 40,000 women in Europe in order to investigate physical, psychological, and sexual violence as well as sexual harassment. The questionnaire recorded frequencies of violence occurrence and distinguished between different kinds of perpetrators (e.g. partner, non-partner). It also asked interviewees to identify the incidents that they felt had had the greatest 'impact' on their life, but did not specifically tailor the questionnaire to obtaining a severity scale suitable for incorporating severity in the measurement of violence.

There are reasons for the latter choice. When we think of severity of violence in our society, especially against women, we refer to how much harm it causes. Violence may cause physical, psychological, economic and moral harm in any combination, and we tend to think of the severity that we attribute to each act of violence as the combined effect. If the act recurs, frequency adds to severity, but the two notions ought to be kept distinct. It is hardly surprising, therefore, that the severity scales which populate the literature try to capture this or that type of harm (Marshall, 1992; Straus *et al.*, 1996). Considerable research has gone into identifying and evaluating the physical, mental and economic consequences of violence against women, and more so recently (EIGE, 2014; Black, 2011; Breiding *et al.*, 2008; WHO, 2013). Research is also starting seriously to investigate moral harm, a much more elusive concept (Mosko, 2011; Koppelman, 2005). However, even if we were willing to concede that consensus on how to measure each type of harm could somehow be reached, aggregating measures of harm across different types may well defy commensurability or be too costly.

Yet, if asked, people are generally able to rank different forms of VAW with respect to severity, implying that they have somehow formed a notion of overall harm. Subjective severity scales exploit this notion (Follingstad 2011; Hudson and McIntosh, 1981; Marshall, 1992; Rodenburg and Fantuzzo, 1993), but they are clearly vulnerable to the criticism of subjectivity. Our intent in this paper is to propose a comprehensive but objective scale which goes some way towards remedying the main shortcomings of existing scales and yet is rather parsimonious in terms of the information it needs. The underlying assumption may be expressed in simple terms. For each main type of violence against women – in this paper we will confine analysis to

physical, sexual, and psychological violence – and within a fairly homogeneous cultural and economic setting, forms of violence that are more widespread in the representative population (i.e. high prevalence violence items) tend to be socially assessed as less severe. We will interchangeably refer to each of these forms of violence as an ‘act’ or an ‘item’. For example, in European countries being pushed or shoved by one’s partner – an act of physical violence – occurs more frequently in the population than being hit by the partner with a hard object – another act of physical violence; and the latter is usually considered to be more severe. Continuing with examples of physical violence, it is hard to dispute that intentional homicide of women (so called *femicide*) is the most severe item on any possible scale. According to Eurostat crime series, recorded prevalence rates for *femicides* are below 1 woman every hundred thousand in the vast majority of European countries and not much above in the remaining minority. Compare, now, intentional killing with a comparatively rare item of (non-lethal) physical violence: being burned. According to the FRA survey data, 140 European women in every hundred thousand reported having been burned at least once by their partner in the year preceding the survey. Leaving aside for the moment problems of comparability across countries and between sources, being burned occurs rarely – it is the least common item of physical violence (from partners) recorded by the FRA survey – and is generally considered serious violence. Fortunately, however, femicide – occurs much more rarely and few would dispute that being killed is more serious than being burned. Examples of sexual and psychological violence tell a similar story. Being belittled or humiliated by one’s partner in private or in public is a form of psychological violence that women in Europe experience fairly frequently according to the quoted survey: almost eighteen thousand in every hundred thousand women report having experienced it during the current intimate relationship. Significantly, it is almost twice as prevalent as being scared or intimidated on purpose by the current partner (for example by yelling and smashing things), which most people would consider more serious. And it is more than five times as prevalent as being threatened to be hurt. The (inverse) prevalence ranking on these acts of violence is indeed consistent with the results of the surveys conducted by Follingstad (2011) and Follingstad and Rogers (2014) on a sample of US adults about perceived severity, emotional and behavioral impact of psychological abuse items.

It may be tempting to dismiss the message that these examples convey as based on a tautology. The argument might run as follows: if an act of violence is common enough, we tend to consider it less severe simply because familiarity makes it less threatening. However, this reasoning is hardly convincing when we look at the higher end of the severity spectrum. To take an extreme example, do we consider femicide more serious than slapping simply because it is a less familiar occurrence? A tautological explanation clearly runs counter to the idea that severity ought to bear some correspondence with the actual harm which is caused.

The core assumption on which this paper is built, in fact is , that the prevalence of a given act of violence in the representative population can be used to (inversely) proxy socially perceived severity. A comprehensive scale can be built upon this premise, which has several advantages. It is objective and yet sufficiently intuitive,

and it is flexible because it can be obtained from different types of surveys. At the same time, it is parsimonious in terms of information because it only requires that the survey be administered to a representative population and that data are gathered on the individual frequency and population prevalence of each item of violence. Finally, it is less susceptible to cultural bias. For example, the widely used Conflict Tactic Scale (CTS), which we shall discuss later, reflects physical harm much more than emotional or psychological harm partly because it is inspired by legal notions about violence that prevail in the USA. The CTS was first developed in the 1970s and was updated in the 1990s. Arguably, however, when it is applied to survey results gathered decades later or outside the USA it is susceptible to cultural bias. The scale that we develop is less susceptible to such bias because it is contextual to the survey results to which it is applied.

To construct the scale, we use fuzzy set theory. In the socio-economic domain the theory was first proposed to analyse and measure relative poverty, but has since been extended to a variety of social phenomena including the effects of marital disruption on well-being, multidimensional quality of life, and multidimensional education mismatch. The parallel between analysis of poverty and analysis of violence against women may help clarify the nature and implications of the scale we are about to propose. In poverty analysis, fuzzy set theory makes it possible to go beyond a dichotomous assessment – being poor vs. not being poor – to view poverty as a matter of degree. It also accounts for the multidimensionality of poverty instead of limiting attention to a single dimension, e.g. monetary poverty. The severity scale that we propose is fuzzy because violence is seen to occur in degrees where ‘degree’ gauges the extent of severity. The scale accommodates multidimensionality by accounting for different acts of violence within each violence type, e.g. being slapped or being hit with a hard object are treated as different ‘dimensions’ of physical violence. A final, albeit subtler, parallel with poverty analysis concerns the role of prevalence in the population. In the fuzzy set theory of poverty, if, say, I cannot afford a car, I am considered poorer in a social setting where most people own a car than in a setting where cars are the privilege of a few. In other words, fuzzy set poverty analysis is built on the premise that the more widespread a given instance of deprivation is, the less this specific deprivation contributes to poverty. It is easy to see the parallel with the premise on which we build our analysis of severity of violence, namely that the more widespread a certain act of violence is, the lower the chances are that it is considered severe. For all these reasons we call our scale ‘fuzzy multidimensional scale of violence against women’, ‘fuzzy violence scale’ for short (FVS).

When poverty is seen as a relative concept, linking prevalence of deprivation with how serious poverty is requires little justification. The question is whether linking seriousness of violence and prevalence is equally justified. We believe it is, and in this paper we offer three pieces of evidence in support. We first conduct an internal validation of our notion of severity by assessing the degree of congruence between women’s subjective evaluation of ‘the most serious event of intimate partner violence’ recorded by the FRA survey and the severity ranking of these events that our scale yields.

We then carry out an external validation by comparing our scale to three scales widely used in the literature. Specifically, we assess the degree of congruence between the severity ranking of comparable acts of violence that obtains from our scale and that yielded by the three scales that we selected for comparison: the Revised Conflict and Tactic Scale (CTS-2); the Severity of Wife Abuse Scale (SWAS) and the Index of Wife Abuse (ISA).

Finally, we exploit our scale to unravel a gender equality puzzle which arises from FRA data (FRA, 2014: p. 31-32) and concerns European countries, namely that some of the more gender equal countries record higher prevalence rates for some types of violence; and conversely for less gender equal countries. This result has come as some surprise and it has been rationalized by arguing that women from less gender equal countries are less likely to report violence because they are less aware of what may constitute violence against them (FRA, 2014: p. 25). While awareness may play a role, we show that the puzzle is resolved once violence is weighted by severity using our scale instead of being measured by simple prevalence. Specifically, our results indicate that more gender equal European countries are associated with comparatively lower levels of violence against women once severity of violence is accounted for.

The paper is organized as follows. Section 2 illustrates the formal construction of the fuzzy violence scale and of fuzzy indexes based on this scale. Section 3 applies the scale to the FRA data, details the ranking of physical, sexual and psychological violence which results from application of the scale, and conducts the internal validation. Section 4 carries out external validation by means of comparison with the CTS-2, The SWAS and the ISA scales. Section 5 addresses the gender equality puzzle and Section 6 concludes.

For convenience, application of the scale, internal validation and external validation are all limited to physical, sexual and psychological violence perpetrated by the partner and reported to have occurred over the last 12 months. In our view, focusing on the last 12 months is necessary in order to ensure representativeness of our measures of violence, since individual recollection of abuses is bound to be much more accurate for the most recent incidents. Limiting the focus to violence by the partner is also warranted by the need to compare our scale with the three selected alternatives, given that the latter were specifically developed to capture the seriousness of partner's abuse. However, the fuzzy scale that we propose can be applied to any type of violence and by any perpetrator, as will be clarified in the sections to follow.

2. Violence against women: constructing a fuzzy scale of severity and fuzzy indexes.

The phenomenon of violence against women is often characterized by dichotomizing the population into who has and who has not experienced a specific act of? violence or any type of violence. This approach suffers from an important limitation, since violence is not merely an incident attributed to an individual in terms of presence or absence; rather, it is an occurrence that manifests itself in different shades and degrees. The

fuzzy approach models violence as something that is experienced in degrees rather than simply occurring or not occurring for individuals in the population.

The fuzzy set approach to social phenomena was first proposed by Cerioli and Zani (1990) and Cheli and Lemmi (1995) in the context of poverty measurement. This methodology has latterly been successfully implemented to monitor the effects of marital disruption on well-being (Aassve *et al.*, 2007), to measure the multidimensional education mismatch (Betti *et al.*, 2011), to measure labour participation (Belhadj, 2015), and to study multidimensional measures of quality of life (Betti *et al.*, 2016).

In a recent contribution, Betti *et al.* (2015) put forward a step-by-step procedure for calculating multidimensional fuzzy measures of social phenomena. Our proposal here is to apply the procedure to violence against women. Our aim is to eventually obtain measures of violence which reflect the occurrence of different acts of violence for the same individual, frequency of each act experienced by the individual, and socially perceived severity of each act.

Different surveys may classify items of violence in different ways and may group them differently into types (sexual, physical, and so on). The procedure we are about to illustrate is independent of the way in which acts and types are worded or classified. However, for the results to be comparable across different classifications certain conditions must hold, specifically the same act should be listed under the same type, and the level of detail of the respective classification should be similar.

The procedure involves the following steps:

1. Identifying violence types
2. For each type of violence identifying the different items to be included in the analysis
3. For each item and each individual, determining a membership function in the $[0, 1]$ interval
4. For each item, determining the severity weight to be used for aggregation across items
5. For each type of violence, aggregating individual membership across items to compute type-specific index scores
6. Aggregating type-specific index scores into an overall index score (optional step).

Steps 1 and 2 clearly depend on the data source. Here we refer to three out the four types of violence investigated by the FRA survey – physical, sexual and psychological – for a total of 30 acts of violence (see Appendix for the complete list). Let h denote types of violence and j items or acts of violence, with $h = 1, 2, \dots, m; j = 1, 2, \dots, k_j$.

Step 3 accounts for frequency of each act of violence experienced by the individual (denoted by i). In order to avoid the too simplistic dichotomisation of presence or absence of a certain act of violence, we construct a membership function for each item and each individual; this ranges in the interval $[0, 1]$, and increases as

violence is experienced more frequently by the individual. Frequency of experience is measured in classes denoted by symbol c and ordered from highest to lowest.

For each item j , the membership function for i is determined as follows:

$$\mu_{j,i} = \frac{1 - F(c_{j,i})}{1 - F(1)} \quad (1)$$

where $c_{j,i}$ is the frequency class of the j -th item for the i -th individual, $F(c_{j,i})$ is the value of the j -th item cumulation function for the i -th individual and $F(1)$ is the population value of the cumulation function for the highest frequency class of the j -th item.

In step 4 weights for aggregation across items are defined to proxy the social evaluation of each item with regard to severity. In multidimensional analysis several weighting methods have been proposed. An earlier survey can be found in Filippone *et al.* (2001) where advantages and drawbacks are discussed in detail. More recently an additional categorization has been put forward (see Guio, 2009, among others) where a distinction is made between consensus and prevalence weighting system. In the present work we have adopted the approach proposed by Betti and Verma (2008) where prevalence weights are ‘corrected’ for correlation. The main reason to combine prevalence and correlation is that, while the prevalence component gives greater importance to items that are least widespread or most widespread depending on the nature of the phenomenon under investigation, the correlation component addresses problems of measurement error and of the redundancy. The resulting system of weighting has, in fact, found applications in the study of phenomena as different as marital disruption (Aassve *et al.*, 2007) and educational mismatch at graduate level (Betti *et al.*, 2011)

Following this approach, the severity scale that we proposed is defined as a set of k weights for each type of violence h combining a prevalence component and a correlation component, as follows:

$$w_{hj} = w_{hj}^a \cdot w_{hj}^b \quad (2)$$

The first weighting component w_{hj}^a is the coefficient of variation of item j within type h in the reference population. As shown in Betti and Verma (2008), the coefficient of variation suitably flattens the inverse prevalence distribution in correspondence with low prevalence phenomena like violence. The second component w_{hj}^b accounts for correlations among items: it gives less weight to items more highly correlated with others (correlation weights), so as to reduce the effect of redundancy and arbitrariness in the choice of original items.

We shall refer to each set of severity weights as our ‘fuzzy scale for violence type h ’, formally FVS(h). Note how this implies that acts of violence can only be ranked by severity if they pertain to the same violence type, while ranking across types is not allowed.

In step 5 multidimensionality comes in as individuals are assigned a type-specific violence index score.

For each individual i , aggregation over a set of items within each h type is given by the weighted mean taken over j items:

$$\mu_{h,i} = \frac{\sum_{j=1}^{k_h} (\mu_{j,i} * w_{hj})}{\sum_{j=1}^{k_h} w_{hj}} \quad (3)$$

In plain language, $\mu_{h,i}$ is the value (score) that the index of fuzzy, multidimensional h -type violence takes for individual i accounting for prevalence (which acts of h -type violence the woman in question actually suffered from, if any), as well as individual frequency and degree of severity of each act. We shall refer to $\mu_{h,i}$ as ‘type-specific Fuzzy Index of Violence’, or FIV(h) for short.

Step 6 aggregates FIV(h) across violence types by simply calculating the unweighted mean over the m types as follows:

$$\mu_i = \frac{\sum_{h=1}^m \mu_{h,i}}{m} \quad (4)$$

where μ_i is the value that the fuzzy multidimensional index of all types of violence takes for each individual. Note that aggregation by means of the simple average implies that we are deliberately abstaining from ranking violence types in terms of severity, i.e., sexual violence *per se* is not necessarily more severe than physical violence, since it all depends on the specific acts that are perpetrated. A stance more radical than ours would be that types of violence are in no way commensurable; hence taking the simple average across violence types makes little sense. This is a legitimate stance which, however, only implies that the fuzzy methodological approach should stop at step 5 and forsake step 6.

3. Estimation and internal validation of the Fuzzy Violence Scale using FRA survey data

Application of steps 1-4 to the entire sample of women in the FRA survey yields the scale recorded in Table 1. For the reasons stated earlier, estimation of the scale was confined to acts of sexual, physical and psychological violence reported to have been perpetrated by partners in the 12 months preceding the survey. For each type of violence, the different acts are listed in decreasing order of severity in the table, with the ranking given by the weights reported in the third column. Weights are normalised so that the average equals 1 within each violence type.

Table 1. Fuzzy Violence Scale: Severity weights by act and type of violence based on FRA survey data

Violence Type	Acts (items of violence)	Severity weight	Two-level Severity category
Psychological violence	<i>How often has your current partner</i>		
	Belittled or humiliated you in private?	0.616	Minor
	Got angry if you speak with another man/woman?	0.666	Minor
	Insisted on knowing where you are in a way that goes beyond general concern?	0.703	Minor
	Belittled or humiliated you in front of other people?	0.739	Minor
	Become suspicious that you are unfaithful?	0.747	Minor
	Done things to scare or intimidate you on purpose, for example by yelling and smashing things	0.754	Minor
	Tried to keep you from seeing your friends?	0.774	Minor
	Prevented you from making decisions about family finances and from shopping independently?	0.854	Minor
	Tried to restrict your contact with your family of birth or relatives?	0.915	Minor
	Threatened to hurt you physically?	1.084	Severe
	Forbidden you to work outside the home?	1.141	Severe
	Threatened to take the children away from you?	1.208	Severe
	Forbidden you to leave the house, takes away your car keys or locks you up?	1.279	Severe
	Threatened to hurt or kill someone else/ someone you care about?	1.355	Severe
	Hurt your children?	1.355	Severe
Threatened to hurt your children?	1.392	Severe	
Made you watch or look at pornographic material against your wishes?	1.418	Severe	
Physical violence	<i>Your current partner has...</i>		
	Pushed you or shoved you?	0.748	Minor
	Slapped you?	0.832	Minor
	Grabbed you or pulled your hair?	0.919	Minor
	Thrown a hard object at you?	0.951	Minor
	Beaten you with a fist or a hard object, or kicked you?	0.985	Minor
	Beaten your head against something?	1.097	Severe
	Tried to suffocate you or strangle you?	1.146	Severe
	Burned you?	1.161	Severe
Cut or stabbed you, or shot at you	1.161	Severe	
Sexual violence	Have you consented to sexual activity because you were afraid of what your current partner might do if you refused?	0.936	Minor
	Made you take part in any form of sexual activity when you did not want to or you were unable to refuse?	0.981	Severe
	Attempted to force you into sexual intercourse by holding you down or hurting you in some way?	1.029	Severe
	Forced you into sexual intercourse by holding you down or hurting you in some way?	1.054	Severe

Source: our estimation using FRA survey microdata (entire sample for 28 European countries). Violence from current partner. Sexual and physical violence refer to episodes occurred in the last 12 months.

Later on we shall need a simplified version of our scale so that we can compare it with and validate it against other popular scales. To this end, we collapsed our extended scale into a two-level scale distinguishing only between minor and severe items of violence. The cut-off point between ‘severe’ and ‘minor’ was set to

coincide with the largest difference in severity weights between two consecutive items of the extended scale, as shown in the last column of Table 1. In what follows we shall refer to this two-level ranking as the ‘collapsed fuzzy multidimensional scale’.

Having obtained the fuzzy severity scale, we sought to validate the scale internally, i.e. by exploiting the fact that the FRA survey asked the respondents about the psychological and physical consequences of the most serious incident of violence. Specifically the respondents were asked to identify the acts of physical or sexual violence inflicted upon them by the partner during what they considered to be the most serious incident they had incurred in the last year¹. Having been informed that ‘most serious incident’ should be understood as the one which had the greatest psychological or physical impact on them, the respondents were also asked to indicate the specific effects of the incident by drawing them from three lists. The first list included nine negative feelings (anger; aggressiveness; shock; fear; shame; embarrassment; guilt; annoyance; other). The second list encompassed nine psychological symptoms (depression, anxiety, panic attacks, loss of self-confidence, feeling vulnerable, difficulty in sleeping, concentration difficulties, difficulties in relationships, other), while the third list indicated eight types of injuries (bruises, scratches, wounds, sprains, burns, fractures, broken bones, broken teeth, concussion or other brain injury, internal injuries, miscarriage, other).

We constructed two measures of congruence between our severity scale and the information provided by respondents, which we call item-based and individual-based measures, respectively. Both measures yielded a fair degree of congruence.

The item-based measure was computed for physical violence,² and it represents the ranking correlation coefficient between the severity weight of a given item in our physical violence scale – say, having been grabbed or pulled by the hair – and the likelihood that this item was indicated to have occurred during the most serious incident. We computed such likelihood by taking the share of women indicating that they had suffered from that act of violence during the worst incident over all the women reporting that they had suffered from the same act in the preceding 12 months. The exercise yielded a ranking correlation coefficient 0.862, significant at 1% percent level.

The individual-based measure of congruence relies on the fuzzy multidimensional violence score rather than, simply, the fuzzy scale, where, however, the former incorporates the latter. Specifically, we computed a Pearson correlation coefficient between (i) the number of psychological or physical effects the respondent selected as a result of the incident, and (ii) the respondent’s FIV average score for physical and sexual

¹ Items describing psychological violence were not included in this question, so that the test of internal validity and reliability of our measure considers only physical and sexual violence. The only item of psychological violence included is ‘threatened to be hurt physically’ which, for the purpose of the validity test, we have treated as a physical violence item since in the questionnaire of the FRA survey it is always reported in the list of items referring to physical abuse.

² The number of items listed for sexual violence is too low for correlation analysis; hence calculations were restricted to physical violence.

violence. Concerning the number of reported effects, we tried different specifications, e.g. first including and then excluding emotional effects on the presumption that the latter might capture merely temporary aftereffects. Calculation of individual FIV average scores for sexual and physical violence involved steps 1-6 of our methodology after having excluded psychological or any other forms of violence from any step.

The correlation coefficient that we obtained varies from 0.562 to 0.568 depending on whether merely emotional repercussions were included or excluded, but it remains consistently significant at 1% percent level. In plain words, what we called 'internal validation' appears to have passed the test.

4. External validation of the FVS: comparison with other scales

The next step in the validation process entails comparison between the severity ranking of our fuzzy scale and the ranking of the scales used for comparison. In what follows we first present the three scales used for validation and then verify sameness of ranking for comparable acts of violence.

4.1. The three scales used for validation

This comparative exercise is not straightforward, since existing scales are very heterogeneous regarding classification of violence into acts and types, reference population, type of perpetrator (intimate partner, family members, others), and scoring criteria. In order to ensure meaningful comparisons, we were compelled to:

- (i) considerably simplify scales and
- (ii) select for comparison only scales with classifications similar to that adopted by the FRA survey.

Simplification meant that the comparison had to be carried out using a minimum common ranking, which, in our case coincides with the two level severity scale (minor-severe) in Table 1 above (last column). Selection of comparable classifications meant, in turn, that we had to limit our choice to three scales out of the ten that we originally considered. The ten scales are briefly described in the Appendix (Table A1). The three surviving scales are the Revised Conflict Tactics Scales (CTS-2), the Severity of Violence Against Women Scale (SVAW) and the Index of Spouse Abuse (ISA), as noted. Key features of these scales are briefly summarized below.

The Conflict Tactic Scale was developed in the 1970s by sociologist Murray Straus to study violence within families (Straus, 1979), and it was subsequently revised as CTS-2 in 1996 (Straus *et al.*, 1996) by adding new items and one more type of violence (sexual coercion). It was translated into many languages and it is one of the tools most frequently used to measure intimate partner violence, despite the many criticisms that it has received (Walby, 2005). A bibliography compiled in 2006 counted 352 references to articles reporting results based on the CTS2 (Straus, 2006). The CTS-2 describes how partners deal with interpersonal conflicts in their

intimate relationship by identifying different types of “conflict tactics” – negotiation, psychological aggression, physical assault, sexual coercion. Each conflict tactics scale differentiates maltreatment behaviors between minor and severe violence based on the presumed harm. According to Straus (2007: 191), the latter distinction ‘is roughly parallel to the legal distinction in the United States between "simple assault" and "aggravated assault"', with "aggravated assault" meaning an ‘attack that is likely to cause grave bodily harm’ (Straus, 1990: 58). Despite the fact that the scale evidently draws from the legal notion of violence that prevails in the USA, it has been used in very different cultural contexts. For our purposes, the CTS-2 allows for the widest and most straightforward comparison since it classifies items between severe and minor and covers the same three types of violence on which we focus, although it gives them different names (physical assault, psychological aggression, and sexual coercion)³.

The Severity of Violence Against Women Scale (SVAWS) was developed by Marshall (1992) to provide a measure of severity more sensitive than earlier scales, including the CTS. The SVAWS is based on interviews first administered to 707 college women and later repeated on a sample of 208 adult women not at college. The interviewees were asked to consider 46 acts of violence and rate them on a 10-point scale gauging “how serious, aggressive, abusive, threatening, and violent "it would be if a man did the act to a woman” and how much psychological or emotional harm they would have expected (Marshall, 1992: 106). Comparison between our scale and the SVAWS only proved feasible for physical and sexual violence and required some adaptation⁴. The SVAWS ranks acts of physical violence using four severity levels (mild, minor, moderate, and severe) which can be easily collapsed into a two-level collapsed severity scale. A more serious problem arises for acts of sexual violence, which the SVAWS does not rank according to severity. However, the women that Marshall interviewed rated the physical and emotional impact of sexual violence alongside physical violence, and these ratings are reported in Marshall (1992). We used these ratings to construct a severity ranking of sexually violent acts, replicating what Marshall had done for acts of physical violence.

The last scale that we used for external validation is the Index of Spouse Abuse (ISA), which was proposed by Hudson and McIntosh (1981) and is still reputed to be a well-validated tool to measure intimate partner violence (Signorelli *et al.*, 2014). This scale, too, is based on the ratings of acts of violence – 30 in all – by a sample of women - 188 graduate and undergraduate students and some faculty members from the School of Social Work and the Department of Psychology at the University of Hawaii. Two indexes of severity of abuse (ranging from 0 to 100) were drawn from the student answers: an ISA-P score for intimate physical partner violence (IPV henceforth) and an ISA-NP score for nonphysical IPV. Each index score reflected the frequency of occurrence of a given act of violence, as well as the subjective rating of the seriousness of the

³ Recall, however, that we chose to focus on sexual, physical and psychological violence perpetrated by the partner also because this facilitates comparison with other scales. In principle, the FVS can be applied to all types of violence and all perpetrators.

⁴ The reason for excluding psychological violence from the comparison is that the SVAWS adopts a rather selective notion of psychological violence by considering “only acts which were directly or indirectly physically abusive” (p. 106).

abuse.⁵ Comparison between our FVS and the ISA is only feasible for psychological violence, which, however, nicely complements the comparison between the FVS and the SVAWS. This limitation is due to the fact that the majority of ISA items for physical violence cannot be made comparable to those in the FRA survey. In order to translate the severity ratings underpinning the SVAWS into the two-level ranking of our collapsed fuzzy scale we used the weights appraising severity of abuse that Hudson and McIntosh reported in their original article (Hudson and McIntosh, 1981: 880). Again, we set the cut-off point between minor and severe items of psychological violence at the largest distance between consecutive items.

4.2 Results of item-based validation

For item-based validation we compute the share of acts of violence that received the same severity ranking in our scale and each of the scales in the comparison (out of the total number of items that we could match, i.e. which we found semantically comparable in the two scales). We assume that a share higher than two thirds may be deemed satisfactory.

Table 2 summarizes the results of our validation exercise by reporting the total number of items matched alongside the number of matched items receiving the same severity ranking in our scale and the one being compared with it⁶.

Table 2. Items matched and congruence of severity ranking: FVS versus CTS-2, SVAWS, and ISA

Scale matched with FVS	CTS-2			VAW		ISA
	Physical	Sexual	Psychological	Physical	Sexual	Psychological
Type of violence	Physical	Sexual	Psychological	Physical	Sexual	Psychological
Number of FMS items matched	9	3	4	9	2	8
Number of FMS items matched with same severity rankings	8	3	4	7	2	6

Taking the share of same ranking matches over total matches as a rough indicator of congruence, congruence with our scale is highest for the CTS-2 (with a share of 94% across types of violence), followed closely by the

⁵ An extended version of the ISA, renamed Partner Abuse Scale, was developed by Hudson in the 1990s in order to take on board newer kinds of partnership that were becoming increasingly common (Attala et al., 1994). We chose however, to consider the original scale, since in the Partner Abuse Scale the severity of single item responses is measured by the frequency and not by the nature of abuse.

⁶ Tables 2A, 3A and 4A in the Appendix list all the items in each of the three scales, distinguishing between those we could match and those we could not match. They also record sameness of ranking (with our scale) on an item-by-item basis, and it put forward possible reasons for discordance where discordance occurs.

SVAWS (81%) and the ISA (75%). In all these three cases the share of congruent same ranking matches is over two thirds.

Summing up on what we termed ‘external validation’, the results show a reassuring degree of congruence between our severity ranking and that of three scales widely used in the literature. Discordance of ranking is the exception rather than the rule. Yet it is important to look closer at some of the exceptions in order to appraise the reasons for discordance. With reference to physical violence, one notable instance of incongruence concerns ‘being beaten with a fist or hard object or being kicked’ which receives a ‘minor’ score in our scale and a ‘severe’ score in both the CTS-2 and the SVAWS. In our view, however, reasons for incongruence are contingent rather than intrinsic. First, in our fuzzy scale this item stands at the border between minor and severe acts of physical violence: with a severity weight of 0.985, it does not score much less than the first ‘severe’ act in the scale, namely when the partner ‘beat your head against something’ (column 3 of Table 1). This means that incongruence is partly the result of having to collapse a nine-step scale into a two-step one. An additional reason for incongruence is the way in which items are identified. ‘Being beaten with a fist or hard object or being kicked’ combines different abusive behaviours. Since our severity weights are inversely proportional to prevalence, by construction collapsing different behaviors into the same category tends to reduce the severity score of the composite category. A similar reasoning applies to the two discordant matches resulting from comparison between the FVS and the ISA scale. Again, the FRA questionnaire combined a range of actions that the ISA questionnaire investigated separately.

5. Revisiting the gender equality puzzle in Europe

In this final section, we show how the fuzzy methodology can be used to improve comparisons across countries with respect to a simpler prevalence-of-violence approach. Specifically, we consider the 28 European countries included in the FRA questionnaire and focus comparison among countries on a topical question, namely whether countries with better track records of gender equality score more favorably also in terms of violence against women. The value added for European or national policy of providing a clear answer to this question should be self-evident.

This is not a new question. On the contrary, it has inspired a large body of research in social sciences since at least the 1970s, when American feminist scholars first rooted violence against women in patriarchal norms and values (Dobash and Dobash, 1979 and Walker, 1989 among many others). To date, however, the answers that have been given to these questions are neither consistent nor conclusive⁷.

⁷ It obviously falls outside the scope of this article to review or even summarize this literature. We can only mention recent and selected contributions with opposing views. For example, FRA (2014) finds a positive correlation between gender equality and

There are substantive reasons for these inconclusive answers. A general reason is that it is difficult to prove causality in this area of research. Take the issue of awareness that was raised earlier in the paper. To the extent that gender equality raises awareness among women of what constitutes violence, it is bound to increase reported prevalence of violence in more gender equal countries. Hence, if we are interested in gauging whether economic independence or better education for women affect violence, any causal effect of these dimensions of gender equality may be confounded by the rise in awareness. An additional reason for inconclusive answers is that violence against women changes form and (female) targets in response to female emancipation (Bettio and Ticci, 2017). This makes it difficult to put a plus or a minus on a phenomenon that evolves.

But there may also be methodological reasons. The specific reason which we emphasize here is that aggregate comparisons across countries generally use prevalence indicators, e.g. the share of women that have been exposed to any act of violence in the last 12 months or during their lifetime. Since such aggregate measures of prevalence do not account for severity, and they often do not even account for frequency, they are likely to yield partial answers at best.

In order to pursue this argument, we proceed as follows. We first compute two distinct sets of country scores for violence by intimate partners. The first set uses a simple prevalence indicator, while the second set uses our FIV score for the female population of that country⁸. We then correlate the two sets of country scores with the 2012 authoritative index of gender equality elaborated by the European Institute for Gender Equality (henceforth Gender Equality Index, GEI for short: EIGE 2016) and compare results. GEI is a composite statistic ranging between 1 (total inequality) and 100 (full equality). It is based on Amartya Sen's capability approach, as well as on widely accepted notions of equality, and it measures gender gaps in six main domains: work, money, knowledge, time, power, health⁹.

The main findings from this exercise is that our fuzzy score of violence correlates negatively and significantly with gender equality across countries, whether we consider physical, sexual and psychological FIV separately or whether we aggregate them into the overall FIV. In contrast, prevalence measures of violence yield mixed results. The details are set out in Figure 1 and Table 3 below. Figure 1 divides European countries into the four quartiles of the 2012 distribution of the GEI Index¹⁰. The average FIV is then computed for the countries

lifetime prevalence of physical and/or sexual violence; Straus (1994) and Aizer (2010) estimate an inverse relation; while Yodanis (2004) and Gressard *et al.* (2015) find either a negative or a non-significant association depending on the type of violence.

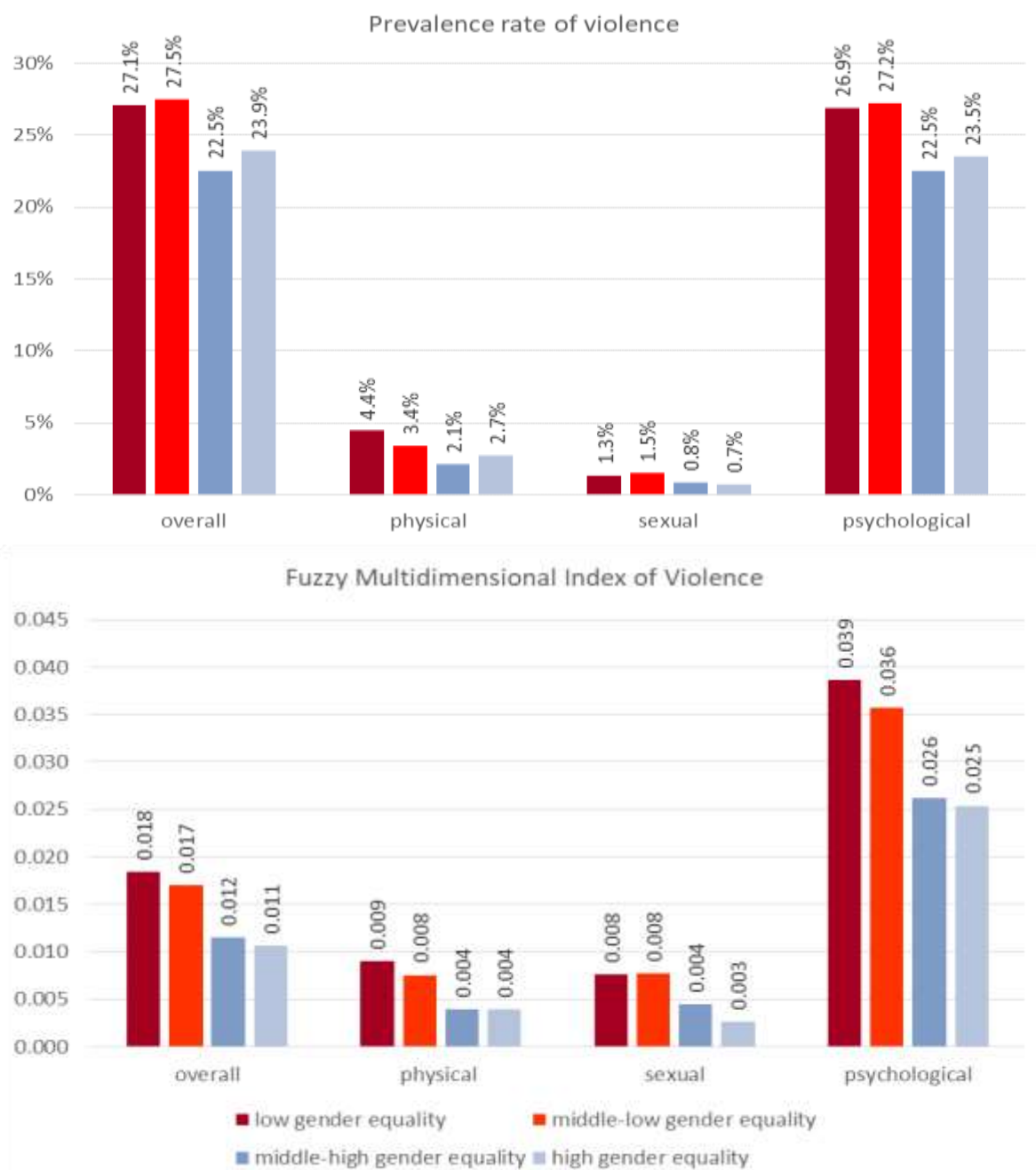
⁸ Here too we consider violence reported to have occurred in the 12 months preceding the interview. The drawback is that the number of women reporting any violence may be small for some countries. This is why we are not providing individual country figures for prevalence rates or the FIV index. However, they are available upon request.

⁹ EIGE publishes also two satellite indices on intersecting inequalities and violence. These indices are conceptually related to gender equality, but are not included in the core index because they measure phenomena affecting a selected group of the population (EIGE, 2013).

¹⁰ The same year as the FRA survey.

in each quartile and for overall, sexual, physical and psychological violence. The FIV score decreases clearly and consistently from the first to the fourth quartile indicating that, if we account for frequency and severity

Figure 1. Average prevalence rates and fuzzy, multidimensional indexes of intimate partner violence for countries in each quartile of the Gender Equality Index distribution.



Source: FRA violence against women survey dataset 2012 and European Institute of Gender Equality (EIGE) accessed in January 2017.

Notes:

- The reference population includes all women who were married, living with an unmarried partner, or involved in a relationship at the time of the interview. Prevalence rates of physical and sexual violence represent the percentage of these women having experienced violence at least once in the 12 months prior to the interview. For psychological violence, the percentage refers to respondents who reported having experienced psychological violence at least sometimes during the current relationship.
- Low gender equality countries include: BG, EL, HR, LT, PT, RO, SK; Low to middle gender equality countries include: CY, CZ, HU, IT, LV, MT, PL; Middle to high gender equality countries include: AT, DE, EE, ES, FR, IE, LU; High gender equality countries include: BE, DK, FI, NL, SE, SI, UK.

besides prevalence, physical, sexual and psychological violence more than doubles when we move from countries with the highest GEI scores to those with the lowest scores.

Prevalence rates show a much less consistent pattern. For instance, prevalence of physical and psychological violence is higher for countries in the top quartile than for those in the third quartile. Moreover, prevalence of sexual violence is higher for countries in the second quartile compared to the first.

Table 3 report signs, values and significance levels for the Pearson and the Spearman correlation coefficients between each European country FIV score and its gender equality index, with calculations repeated for overall violence and for the three main types. All the coefficients turn out to be negative; all except one are above 50%: and all are statistically significant at the conventional level or lower. When prevalence rates are used instead of FIV scores, the negative sign is there but values and significance drop considerably: for example, only three coefficients out of eight show levels of significance above the conventional 5%.

Table 3. Correlation between indicators of violence and the 2012 EIGE index of Gender Equality.

	FIV				Prevalence			
	<i>OVERALL IPV</i>	<i>physical IPV</i>	<i>sexual IPV</i>	<i>psychological IPV</i>	<i>OVERALL IPV</i>	<i>physical IPV</i>	<i>sexual IPV</i>	<i>psychological IPV</i>
Pearson correlation								
2012 EIGE Gender Equality Index	-0.581	-0.541	-0.522	-0.506	-0.167	-0.432	-0.381	-0.178
	<i>0.001</i>	<i>0.003</i>	<i>0.004</i>	<i>0.006</i>	<i>0.397</i>	<i>0.022</i>	<i>0.046</i>	<i>0.366</i>
Spearman correlation								
2012 EIGE Gender Equality Index	-0.591	-0.551	0.495	-0.521	-0.200	-0.499	-0.327	-0.195
	<i>0.001</i>	<i>0.002</i>	<i>0.007</i>	<i>0.005</i>	<i>0.308</i>	<i>0.007</i>	<i>0.089</i>	<i>0.319</i>

Source: FRA violence against women survey dataset 2012, and European Institute of Gender Equality (EIGE) accessed in January 2017.

Notes: the p-value/significance level of each correlation coefficient is in italics. Correlation coefficients with significance at 5% level or lower are shown in bold. See Table 3 for details on the reference population and definition of prevalence rates.

Taken together, the above results suggest that violence by the intimate partner tends to be lower in more gender equal countries, provided that severity and frequency of abuse are taken into account. Of course, correlation does not imply causality, but proving causality was not the purpose of our exercise.

6. Summary discussion

In this paper we have taken a fresh look at how violence against women is measured. Our specific goal has been to develop a scale of severity of violence against women able to remedy important shortcomings of many existing scales, primarily subjectivity and the risk of transposing perceptions of violence that are time- and culture-bound to unwarranted times and places. In order to build our scale we resorted to fuzzy set theory and used it to operationalize our key assumption that, in a given cultural setting and for each type of violence, inverse prevalence proxies the social evaluation of the severity attached to a given abusive act. The scale that we developed and called 'fuzzy scale of violence' can be used to construct individual, area-specific, type-specific but also overall indexes of violence that account at the same time for prevalence, frequency and severity of violence.

Validation of our scale against three well known alternatives – the revised Conflict Tactic Scale, the Severity of Violence Against Women Scale and the Index of Wife Abuse – revealed a high degree of congruence of the severity ranking with our FVS. Given strong congruence, the question naturally arises as to whether the scale we propose is redundant. We see four reasons to discount this possibility. First, our scale (and the related indexes) are much less 'specialized' than most existing alternatives. The latter target intimate partner violence to the exclusion of violence from non-partners, and sexual harassment in particular. For the purpose of comparison, we have also chosen to confine our calculations to intimate partner violence in this paper. However, the scale and indexes we propose can be easily extended to all types of violence and by any perpetrator. Moreover, our fuzzy scale is less costly to implement because it can be easily derived from any representative and sufficiently well-articulated survey of violence, whereas use of other scales may require the survey questionnaire to be tailor-made to match the chosen scale. Third, the methodology which underpins our scale is not context specific, although the results are. In other words, our methodology can be applied in the same way to the results of a survey of violence in, say, a small village in Asia and a large metropolitan city in the rich West. The resulting severity scales or indexes will differ to the extent that the social evaluation of violence differs in these two social contexts; but the risk will be avoided of transposing to the Asian village the subjective evaluation of violence expressed by the (female) population of a large Western city.

Finally, we have shown that our violence metrics can improve comparisons among countries compared to the most widely used metrics in comparative analysis, namely prevalence rates. In order to illustrate improvement we have shown that the country violence scores that our fuzzy approach yields reveal an inverse association between IPV violence against women in a given country and the degree of gender equality in that country. Although this result is rather suggestive and appears to resolve a puzzle which has been debated in the literature, there are two important reasons for caution. The first is that we only verified association, not causation. The second is that what holds for intimate partner violence may not hold for

violence by non-partners, and for sexual harassment in particular. In fact, sexual harassment tends to be higher for women who work, and it is therefore likely to increase rather than decrease as women reduce their gaps with men (Bettio and Ticci, 2017). However, these limitations are not intrinsic to the metrics that we propose; rather, they simply invite additional research.

References

- Aassve A., Betti G., Mazzucco S. and Mencarini L. (2007), Marital Disruption and Economic Well-being; A Comparative Analysis, *Journal of the Royal Statistical Society, Series A*, **170**, pp. 781–799.
- Aizer, A. (2010), The Gender Wage Gap and Domestic Violence. *American Economic Review*, 100(4), pp. 1847-1859.
- Attala J.M., Hudson, W.W, McSweeney, M., Haworth Continuing Features Submission (1994), A Partial Validation of Two Short-Form Partner Abuse Scales, *Women & Health*, 21(2-3): 125-139.
- Belhadj B. (2015), Employment measure in development countries via minimum wage and poverty: new Fuzzy approach, *Opsearch*, 52, pp. 329–339.
- Betti G., D’Agostino A. and Neri L. (2011), Educational Mismatch of Graduates: a Multidimensional and Fuzzy Indicator, *Social Indicators Research*, 103, pp. 465–480.
- Betti G., Gagliardi F., Lemmi A. and Verma V., Comparative Measures of Multidimensional Deprivation in the European Union, *Empirical Economics*, 49 (2015), pp. 1071–1100.
- Betti G., Soldi R. and Talev I. (2016), Fuzzy Multidimensional Indicators of Quality of Life: The Empirical Case of Macedonia, *Social Indicators Research*, 172, 39–53.
- Betti G., Verma V. (2008), Fuzzy measures of the incidence of relative poverty and deprivation: a multi-dimensional perspective, *Statistical Methods and Applications*, 12(2), pp. 225-250.
- Bettio, F. and Ticci, E. (2017), *Violence Against Women and Economic Independence*, Luxembourg: Publications Office of the European Union
- Black, M.C. (2011). Intimate partner violence and adverse health consequences: Implications for clinicians. *American Journal of Lifestyle Medicine*, 5, 428–439.
- Breiding, M.J., Black, M.C., & Ryan, G.W. (2008). Chronic disease and health risk behaviors associated with intimate partner violence—18 U.S. states/territories, 2005. *Annals of Epidemiology*, 18, 538–544.
- Cerioli A. and Zani S. (1990), A fuzzy approach to the measurement of poverty, in *Income and wealth distribution, inequality and poverty*, eds. C. Dagum and M. Zenga (1990), pp. 272–284, Springer Verlag, Berlin.
- Cheli B. and Lemmi A. (1995), A Totally Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty, *Economic Notes*, **24**, pp. 115–134.
- Dobash R.E., Dobash R. (1979), *Violence Against Wives: A Case Against the Patriarchy*. New York: Free Press
- EIGE – European Institute for Gender Equality (2013). Gender Equality Index Report, available at <http://eige.europa.eu/sites/default/files/documents/Gender-Equality-Index-Report.pdf>.
- EIGE (2016), Gender Equality Index 2015 Measuring gender equality in the European Union 2005–2012, Luxembourg: Publications Office of the European Union.
- European Institute for Gender Equality, 2014. *Estimating the costs of gender-based violence in the European Union Report*, Luxembourg: Publications Office of the European Union, 2014.

Filippone A, Cheli B, D'Agostino A (2001) Addressing the interpretation and the aggregation problems in totally fuzzy and relative poverty measures. ISER Working Paper Series number 2001–22, University of Essex.

FRA – European Union Agency for Fundamental Rights (2014). Violence against women: an EU-wide survey. Main results, Vienna.

Follingstad, D. R. (2011), A measure of severe psychological abuse normed on a nationally representative sample of adults. *Journal of Interpersonal Violence*, 26: 1194–1214.

Follingstad, D.R. and Rogers, M.J. (2014), The Nature and Prevalence of Partner Psychological Abuse in a National Sample of Adults, *Violence and Victims*, 29(1): 3-23.

Gressard L.A., Swahn M.H., Tharp A.T. (2015). A First Look at Gender Inequality as a Societal Risk Factor for Dating Violence. *American Journal of Preventive Medicine*. 49(3): 448-57.

Guio, A-C. (2009), What can be learned from deprivation indicators in Europe?, Eurostat methodologies and working paper, 40 p. Eurostat, Luxembourg.

Hegarty K., Sheehan, M. and Schonfeld, C. A (1999), A multidimensional definition of partner abuse: development and preliminary validation of the Composite Abuse Scale, *Journal of Family Violence*, 14: 399–415.

Hudson, W.W. (1997), *The WALMYR assessment scales scoring manual*. Tallahassee (FL): WALMYR Publishing Company.

Hudson, W.W. and McIntosh, S.R. (1981), The Assessment of Spouse Abuse: Two Quantifiable Dimensions, *Journal of Marriage and Family*, 43(4): 873-885.

Koppelman A. (2005). Does Obscenity cause moral harm? *Columbia Law Review*, 105: 1635—1679

Mosko M. (2011). Sexualized Violence, Moral Disintegration and Ethical Advocacy, Marquette University: e-Publications@Marquette

Marshall, L.L. (1992), Development of the Severity of Violence Against Women Scales, *Journal of Family Violence*, 7(2): 103-121.

Murphy, C.M. and Hoover, S.A. (1999), Measuring emotional abuse in dating relationships as a multifactorial construct. *Violence and Victims*, 14: 39-53.

Rodenburg, F.A., Fantuzzo J.W. (1993), The Measure of Wife Abuse: steps toward the development of a comprehensive assessment technique. *Journal of Family Violence*; 8: 203–228.

Shepard M.F. and Campbell J.A. (1992), The Abusive Behavior Inventory: a measure of psychological and physical abuse. *Journal of Interpersonal Violence*, 7: 291–305.

Signorelli, M.S., Arcidiacono E., Musumeci G., Di Nuovo, S., Aguglia, E. (2014) Detecting Domestic Violence: Italian Validation of Revised Conflict Tactics Scale (CTS-2), *Journal of Family Violence*, 29:361–369

Straus, M. A. (1979), Measuring Intrafamily Conflict and Violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and Family*, 41(1): 75-88.

- Straus, M. A. (1990). The Conflict Tactics Scale and its critics: An evaluation and new data on validity and reliability. In M. A. Straus & R. J. Gelles, *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 49-73). New Brunswick, NJ: Transaction Publishing.
- Straus, M. A. (1994). State-to-state differences in social inequality and social bonds in relation to assaults on wives in the United States. *Journal of Comparative Family Studies*, 25, 7-24.
- Straus, M. A. (2006), Bibliography of the Revised Conflict Tactics Scales (CTS2) and the Parent-Child Conflict Tactic Scales (CTSPC), in Straus, Murray, *Conflict Tactics Scales (CTS) Sourcebook*. Durham, NH: University of New Hampshire, Family Research Laboratory.
- Straus, M. A. (2007). "Conflict Tactics Scales." pp. 190 - 197 in *Encyclopedia of Domestic Violence*, N. A. Jackson. New York: Routledge: Taylor & Francis Group.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S. and Sugarman, D. B. (1996), The Revised Conflict Tactics Scales (CTS2): Development and Preliminary Psychometrics Data, *Journal of Family Issues*, 17(3), 283-316.
- Tolman, R.M. (1989), The development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims*, 4:159–177.
- Walby, S. (2005), Improving the statistics on violence against women, *Statistical Journal of the United Nations ECE*, 22: 193–216.
- Walker, L. E. (1989). Psychology and violence against women. *American Psychologist*, 44(4), 695-702.
- WHO, London School of Hygiene and Tropical Medicine. *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.
- Yodanis C.L. (2004). Gender inequality, violence against women, and fear: a cross-national test of the feminist theory of violence against women. *Journal of Interpersonal Violence*, 19(6):655-75.

APPENDIX

Table 1A. Scales considered for validation and scales eventually used.

<i>Scale and authors</i>	<i>Used for validation</i>	<i>Reason for not using or for selective use</i>
REVISED CONFLICT TACTICS SCALES (CTS-2), Straus et al. (1996)	Yes	
SEVERITY OF VIOLENCE AGAINST WOMEN SCALE (SVAW), Marshall (1992)	Selectively, only for physical violence and sexual violence	Psychological violence is not included in the scale. This scale includes only acts that involve some amount of physical threat.
INDEX OF SPOUSE ABUSE (ISA), Hudson and McIntosh (1981)	Selectively, only for psychological violence	Several items of physical violence are not comparable with the items in FRA survey.
MEASURE OF WIFE ABUSE (MWA), Rodenburg and Fantuzzo (1993)	No	Items are not classified according to their severity
ABUSIVE BEHAVIOR INVENTORY (ABI), Shepard and Campbell (1992)	No	Items are not classified according to their severity
MULTIDIMENSIONAL MEASURE OF EMOTIONAL ABUSE (MMEA), Murphy and Hoover (1999).	No	Items are not classified according to their severity
PHYSICAL (PASP) AND NON-PHYSICAL (PASNP) PARTNER ABUSE SCALE, Hudson (1997)	No	Items are not classified according to their severity
PSYCHOLOGICAL MALTREATMENT OF WOMEN INVENTORY (PMWI), Tolman (1989)	No	Items are not classified according to their severity
COMPOSITE ABUSE SCALE (CAS), Hegarty et al. (1999).	No	The rationale behind severity assessment is different. Severe violence, represented by the category “severe combined abuse” (SCA) includes a combination of physical, sexual, and emotional abuse. SCA is separated from other three dimensions: emotional abuse, physical abuse, and harassment. The authors stress the importance of combining different types of violence and their aim: SCA is defined as “moderate to severe intermittent physical and/or sexual abuse associated with regular emotional abuse, which occurs as part of coercive or control tactics used mainly by male partners against women” (Hegarty et al. 1999: p. 413).
MEASURE OF PSYCHOLOGICALLY ABUSIVE BEHAVIORS (MPAB), Follingstad (2011)	No	The rationale behind severity assessment is different. Severity of violence does not vary across items (as in FRA survey) but for each item, three levels of severity are identified.

Table 2A. Matched and non-matched items in the scales used for validation: physical violence
(orange cells indicate non comparability; red cells indicate incongruence of severity grading with our scale)

<i>FRA Physical violence</i> (...how often has something like this happened to you? Your current partner has...)	REVISED CONFLICT TACTICS SCALES (CTS-2) <i>Physical Assault Scale</i> How often did this happen? How often did your partner do this to you?	SEVERITY OF VIOLENCE AGAINST WOMEN SCALE (SVAW)*† How often has your partner...	COMMENTS
Minor - pushed you or shoved you?	Minor - Pushed or shoved you	Mild - Pushed or shoved you	Ok

Minor - slapped you?	Minor - Slapped you	Moderate - Slapped you with the palm of his hand Moderate - Slapped you with the back of his hand Moderate - Slapped you around your face and head	Ok, not inconsistent
Minor - thrown a hard object at you?	Minor - Thrown something at you that could hurt	Serious - Hit you with an object	Congruence between FRA and CTS-2. Incongruence between FRA and SVAW but the different meaning of the verbs (to hit versus to throw) can be an explanation.
Minor - grabbed you or pulled your hair?	Minor - Twisted your arm or hair Minor - Grabbed you	Mild - Grabbed you suddenly or forcefully Minor - Pulled your hair	Ok
Minor - beaten you with a fist or a hard object, or kicked you?	Severe - Punched or hit you with something that could hurt Severe – Beaten you Severe - Kicked you	Serious - Kicked you Serious - Stomped on you Serious - Used a club-like object on you	Incongruence between FRA and both CTS2 and SVAW. 1) The FRA item includes a larger number of actions. 2) By disaggregating it in different item in line with other scales we expect more congruence. 3) Severity weight of this item is borderline, namely close to that of severe items.
Severe - Burned you?	Severe - burned or scalded you on purpose	Serious - Burned you with something	Ok
Severe - tried to suffocate you or strangle you?	Severe - Choked you	Serious - Choked you	Ok
Severe - cut or stabbed you, or shot at you?	Severe - Used a knife or gun on you	Serious - Used a knife or gun on you.	Ok
Severe - beaten your head against something?	Severe Slammed you against a wall	Serious - Beaten you up	Ok
		Mild violence - Held you down, pinning you in place Mild violence - Shaken or roughly handled you Minor violence - Scratched you Minor violence -Twisted your arm Minor violence –Spanked you Minor violence -Bitten you Serious violence - Punched you.	Not comparable

Note: * We refer to the classification of the items used for this scale rather than to an ordering based on the impact weights proposed by Marshall (1992), since weights can be ordered according to two criteria, namely the physical harm or the emotional harm associated with each item by the women interviewed in the study. Given the relatively low correlation between the two types of harm, the resulting ordering would be different. However, it is worth observing that the classification among mild, minor, moderate and serious violence is in line with the scale based on the impact weighting of the expected physical harm.

† SVAW also includes other types of item for Symbolic violence, Threats of mild violence, Threats of moderate violence, and Threats of serious violence.

Table 3A. Matched and non-matched items in the scales used for validation: sexual violence
(orange cells indicate non comparability; red cells indicate incongruence of severity grading with our scale)

FRA Sexual violence <i>How often has your current partner done any of the following to you?</i>	REVISED CONFLICT TACTICS SCALES (CTS-2) <i>Sexual Coercion Scale</i> <i>How often did this happen? How often your partner did this to you?</i>	SEVERITY OF VIOLENCE AGAINST WOMEN SCALE* <i>How often has your partner...</i>	COMMENTS
Severe - forced you into sexual intercourse by holding you down or hurting you in some way?	Severe - Used force (like hitting, holding down, or using a weapon) to make you have oral or anal sex.	Severe: Physically forced you to have sex	Ok

	Severe - Used force (like hitting, holding down, or using a weapon) to make you have sex		
Severe - apart from this, attempted to force you into sexual intercourse by holding you down or hurting you in some way?			
Severe - apart from this, made you take part in any form of sexual activity when you did not want to or were unable to refuse?	Severe - Used threats to make you have oral or anal sex Severe - Used threats to make you have sex	Severe - Made you have oral sex against your will Severe - Made you have sexual intercourse against your will Severe - Made you have anal sex against your will	Ok
Minor - have you consented to sexual activity because you were afraid of what might happen if you refused?	Minor - Insisted on sex when you did not want to (but did not use physical force). Minor - Insisted you have oral or anal sex (but did not use physical force).		
	Minor - Made you have sex without a condom		Not comparable
		Minor: Demanded sex whether you wanted it or not	Not comparable
		Severe: Used an object on you in a sexual way	Not comparable

*SVAW Scale does not distinguish different categories of sexual violence severity, but we can create them by using the impact weights estimated in the study and creating a divide between minor and severe sexual violence at the biggest distance between consecutive items. According to this method, weights for both physical harm and for emotional harm of the SVAW Scale identify as acts of minor sexual violence the item “Demanded sex whether she wanted to or not”, while the remaining acts can be defined as acts of severe sexual violence (“Made her have oral sex against her will”, “Made her have sexual intercourse against her will”, “Made her have anal sex against her will”, “Used an object on her in a sexual way”)

Table 4A. Matched and non-matched items in the scales used for validation: psychological violence
(orange cells indicate non comparability; red cells indicate incongruence of severity grading with our scale)

FRA Psychological violence (How often does your current partner...)	REVISED CONFLICT TACTICS SCALES (CTS-2) <i>Psychological Aggression Scale</i> How often did this happen? How often has your partner done this to you?	INDEX OF SPOUSE ABUSE (ISA) ‡ Answer each item as carefully and accurately as you can by placing a number beside each one as follows	COMMENTS
Minor - try to keep you from seeing your friends?		Minor - My partner does not want me to socialize with my female friends	Ok
Minor - try to restrict your contact with your family of birth or relatives?			
Minor - insist on knowing where you are in a way that goes beyond general concern?			
Minor - get angry if you speak with another man? (or another woman, if the partner is a woman)		Minor - My partner is jealous and suspicious of my friends	Ok
Minor - become suspicious that you are unfaithful?		Minor - My partner is jealous and suspicious of my friends	Ok

Minor - prevent you from making decisions about family finances and from shopping independently?		My partner is stingy in giving me enough money to run our home	Ok
Severe - forbid you to work outside the home?			
Severe - forbid you to leave the house, take away car keys or lock you up?			
Minor - belittled or humiliated you in front of other people?	Minor - Insulted or swore at you Minor - Said something to spite you	Minor- My partner belittles me Minor - My partner belittles me intellectually	Ok
Minor - belittled or humiliated you in private?	Minor - Insulted or swore at you Minor - Said something to spite you	Minor - My partner belittles me Minor - My partner belittles me intellectually Minor - My partner tells me I really couldn't manage or take care of myself without him	Ok
		Severe - My partner tells me I am ugly and unattractive	Congruence between FRA and CTS-2. Incongruence between FRA and ISA. A possible explanation is that the FRA item includes a larger number of actions.
Minor - done things to scare or intimidate you on purpose, for example by yelling and smashing things?	Minor - Shouted or yelled at you Minor - Stomped out of the room or house or yard during a disagreement	Severe: My partner screams and yells at me	Congruence between FRA and CTS-2. Incongruence between FRA and ISA. A possible explanation is that the FRA item includes a larger number of actions.
Severe - threatened to take the children away from you?			
Severe - hurt your children?			
Severe - threatened to hurt or kill someone else you care about?			
Severe - threatened to hurt you physically?	Severe - Threatened to hit or throw something at you	Severe: My partner frightens me	Ok
	Severe - Destroyed something belonging to you		Not comparable
		Minor - My partner is not a kind person	Not comparable
		Minor - My partner becomes very upset if dinner, housework or laundry is not done when he thinks it should be	Not comparable
		Minor - My partner becomes very angry if I disagree with his point of view	Not comparable
		Minor - My partner becomes surly and angry if I tell him he is drinking too much	Not comparable
		Minor - My partner demands obedience to his whims	Not comparable

		Minor - My partner acts like I am his personal servant	Not comparable
		Minor - My partner feels that I should not work or go to school	Not comparable
		Minor: My partner demands that I stay home to take care of the children	Not comparable
		Severe - My partner orders me around	Not comparable
		Severe - My partner treats me like a dunce	Not comparable
		Severe - My partner has no respect for my feelings	Not comparable
		Severe- My partner insults or shames me in front of others	Not comparable